THE JOURNAL OF PARAPSYCHOLOGY

Volume 18 SEPTEMBER, 1954 Number 3

PAGE

Picture-Frustration Ratings and ESP Scores for Subjects Who Showed Moderate Annoyance at the ESP Task 137
GERTRUDE R. SCHMEIDLER

Clock Cards and Differential Scoring Techniques 153
A. M. J. MITCHELL AND G. W. FISK

A Review of the Pearce-Pratt Distance Series of ESP Tests 165
J. B. RHINE AND J. G. PRATT

A Reappraisal of the Mathematical Evaluation of the Reinforcement Effect 178
T. N. E. GREVILLE

Editorial:

Rational Acceptability of the Case for Psi 184

Letters and Notes 195

Book Reviews

Apparitions by G. N. M. Tyrrell 199
L. E. RHINE

Psychical Research Today by Donald J. West 203
R. H. THOULESS

A SCIENTIFIC QUARTERLY DEALING WITH EXTRASENSORY PERCEPTION, THE PSYCHOKINETIC EFFECT, AND RELATED TOPICS

DUKE UNIVERSITY PRESS
DURHAM, N. C.

The Journal *of* Parapsychology

EDITORIAL STAFF

J. B. RHINE and J. G. PRATT, *Editors*
ELIZABETH A. McMAHAN, *Assistant Editor*
I. A. GREENWOOD and T. N. E. GREVILLE, *Statistical Editors*
DOROTHY H. POPE, *Managing Editor*

ADVISORY EDITORIAL BOARD

HORNELL HART
Duke University

H. H. PRICE
University of Oxford

GARDNER MURPHY
Menninger Foundation

ROBERT H. THOULESS
University of Cambridge

This JOURNAL is published on the fifteenth day of March, June, September, and December by Duke University Press, Durham, North Carolina.

Contributions submitted for publication and all editorial communications should be addressed to the Managing Editor, Dorothy H. Pope, College Station, Durham, North Carolina. Correspondence with the editors is advised before submitting articles other than reports of experimentation. Since it does not forward manuscripts by registered mail, the JOURNAL cannot guarantee that they will not be lost in transit, and contributors are urged to keep copies of their papers. All contributions should be typewritten, double spaced. References should be given in the form adopted by the JOURNAL.

Reprints may be ordered when the proof is returned. The JOURNAL will bear one half the cost of reprints up to two hundred copies. Correspondence concerning subscriptions, change of address, back numbers, and other business communications should be addressed to Dorothy H. Pope, College Station, Durham, N. C.

The subscription price is $6.00 a year; single current numbers $1.50. The rate for back volumes is $7.00; for single numbers $1.75. Missing numbers will be supplied free when lost in the mails if written notice is given within one month of the date of issue. All remittances should be made payable to the JOURNAL OF PARAPSYCHOLOGY.

Subscriptions may be obtained in Great Britain through the Cambridge University Press, Bentley House, Euston Road, London, N. W. 1.

Entered as second class matter May 21, 1942, at the post office at Durham, N. C., under the act of March 3, 1879.

DUKE UNIVERSITY PRESS

The Journal *of*
Parapsychology

Volume 18 SEPTEMBER, 1954 Number 3

PICTURE-FRUSTRATION RATINGS AND ESP SCORES FOR SUBJECTS WHO SHOWED MODERATE ANNOYANCE AT THE ESP TASK[1]

By GERTRUDE R. SCHMEIDLER

ABSTRACT: In previous experiments negative correlations were found between subjects' ESP scores and their extrapunitiveness (that is, their tendency to turn their aggressiveness outward) as measured by the Rosenzweig Picture-Frustration Study. Positive correlations were found between ESP scores and impunitiveness (that is, the tendency to evade or avoid any aggressiveness). The results of the research suggested that this relationship applied only to subjects who were frustrated, but not intensely frustrated, by the ESP test.

In the present study, therefore, a method was worked out for selecting those subjects who were somewhat annoyed, but not intensely so, by the ESP test. The ESP scores of this group (as was predicted) showed a negative correlation with extrapunitive responses and a positive correlation with impunitive responses. These correlations were significantly different from chance.—Ed.

INTRODUCTION

FROM PREVIOUS RESEARCH on ESP the author has concluded, as have many other investigators, that higher ESP scores are likely to be made by subjects with friendly, favorable, interested, and relaxed attitudes than by subjects with critical, hostile, negativistic, and aggressive ones. My own research along these lines started with some experiments in which subjects were divided into "sheep" (those who accepted the possibility of ESP under the conditions of the experiment) and "goats" (those who rejected this possibility) (8). The average score of the sheep for each of six series was above

[1] This research was financed by the Richard Hodgson Fund of Harvard University through the kindness of Dr. Gardner Murphy, the Hodgson Fellow. The author wishes to express her gratitude to him for this and also for his sound critical advice.

chance; that of the goats was below (as if to prove their point that ESP is impossible); and the difference between the two groups was highly significant when the six series were pooled.

In later research, the subjects were given Rorschach tests to determine which were well adjusted and which were poorly adjusted. These subjects, as a group, followed the previous trend. The average score of the sheep was above chance, that of the goats was below, and the difference between them was significant (**6**). Of the four groups (well adjusted sheep, poorly adjusted sheep, well adjusted goats, and poorly adjusted goats) the well adjusted sheep scored the highest and the well adjusted goats the lowest. The difference between these two groups was highly significant. The difference between poorly adjusted sheep and poorly adjusted goats, however, was negligible; both scored near the chance level.

Later two experiments were carried out (**1, 7**) in which the ESP subjects were given the Rosenzweig Picture-Frustration test, a projective test of personality which tries to evaluate how a subject will respond to frustration. The ESP subjects were rated by means of the Rosenzweig test for three traits, described as follows by Rosenzweig (**3**):

"Extrapunitiveness: Aggression is employed overtly and directed toward the personal or impersonal environment in the form of emphasizing the extent of the frustrating situation, blaming an outside agency for the frustration, or placing some other person under obligation to solve the problem in hand.

"Intropunitiveness: Aggression is employed overtly, but directed by the subject against himself in the form of a martyr-like acceptance of the frustration as beneficial, acknowledgment of guilt or shame, or an assumption of responsibility for correcting the frustrating situation.

"Impunitiveness: Aggression is evaded or avoided in any overt form, and the frustrating situation is described as insignificant, as no one's fault, or as likely to be ameliorated by just waiting or conforming."

The test consists of a booklet with 24 cartoons. Each cartoon shows a person in an unpleasant, frustrating situation. Someone is speaking to the frustrated individual, and the remark is printed in a balloon over the speaker's head. Another balloon, but an empty

one, is over the head of the frustrated person. The subject's task is to write in the balloon what the frustrated person might say, and the instructions are to write the first response that comes to mind.

In half the pictures the frustrated person is a man, and in half a woman. The drawings are deliberately crude. Facial features are either indistinct or missing, and the background is sketched with a minimum of detail. The situations themselves are designed to be commonplace, the sort of thing that could happen to almost anyone: mud is splashed by a passing automobile, or a person is not invited to a party. Since almost every subject will be exasperated by contemplating some of the situations, responses to the test may indeed represent behavior when annoyed.

Each response is scored separately for each of the three tendencies (entrapunitiveness, intropunitiveness, and impunitiveness) and each response is given a total score of 1. If the subject has written, for example, "That's your fault. You ought to be more careful," the extrapunitive score is 1. But if he wrote, "We were both at fault," he is given an extrapunitive score of 1/2 and an intropunitive score of 1/2. The total of the three categories is 24. By this system of scoring, a subject who is very high for one tendency must be correspondingly low for one or both of the others.

It was found in the first study, where subjects were tested individually, that those who were the most impunitive and intropunitive gave higher ESP scores than those who were the least impunitive and intropunitive. Those who were the most extrapunitive gave ESP scores that were lower than the scores of the least extrapunitive. These differences were suggestive, but not significant (1).

In a subsequent study, with classroom tests, a low but significant positive correlation was found between ESP score and impunitive responses, and a low but significant negative correlation between ESP score and extrapunitive responses (7).

More detailed analysis of the data suggested, however, that these relationships held only for groups where the ESP test was experienced as somewhat frustrating. It is this observation which the present experiment was designed to test.

The need for further examination of the subjects' attitudes arises from two limitations of the Picture-Frustration Study. In the first

place, it does not attempt to analyze how the subject will respond when he is not annoyed. In the second place, it is only a paper and pencil test, often administered (as in these experiments) in a friendly atmosphere to students who regard taking a projective test as an interesting experience. Any frustration which a subject feels in this setting is likely to be pallid compared with the intensity with which he responds to real life. The vividness of his emotion will probably be lessened further by the fact that some of the pictures represent scenes in which he has never participated, and that in half of the pictured situations he is asked to identify with someone of the opposite sex. Thus by the nature of the test we would not expect the subject's scores to show how he would behave in an extremely frustrating experiment, nor in one which was agreeable and interesting to him. We would expect the scores to indicate only how he would behave in an experiment that was basically agreeable or unimportant, but that had some annoying components. It was decided, therefore, that along with the Rosenzweig test some other test should be used in order to show to what extent the subject was frustrated by the ESP test situation.

The procedure of the present experiment, taking this reasoning into account, consisted of administering three tests to the subjects. The first was an ESP test; the second was the Picture-Frustration Study; and the third was a group of measures for rating the subject's frustration during the ESP runs. The hypothesis was that subjects who were rated as moderately annoyed by the ESP test would have a clearer relationship between ESP scores and Picture-Frustration scores than the total group; that these moderately annoyed subjects would show a positive relationship between ESP scores and impunitive scores on the Picture-Frustration Study; and that the same group of subjects would show an inverse relationship between ESP scores and extrapunitive scores.

Procedure

ESP Test

The method of administering the ESP test was described in detail in the author's earlier report (7). It will therefore be summarized only briefly here.

The experiment was conducted in the classroom as part of a course in introductory psychology. The experimenter was the instructor. Stimulus material, or targets, consisted as usual of lists made up of five items arranged in random order according to a table of random numbers. The subjects were instructed to try to guess the order of the items. The lists were kept in opaque envelopes, and were concealed from the subjects until the experiment was completed and the records were in the hands of the experimenter. The experimenter did not know the order of the items at the time of giving the instructions to the subjects, and therefore could not have given clues as to the correct responses. All responses were given two independent scorings to minimize checking errors. The subjects were required to state, before they knew which of their responses were correct, whether they thought the correct responses *might* be due to extrasensory perception, or whether they felt certain that any correct responses they made would be due to chance alone. Those who made the former choice have been referred to as sheep; the others were characterized as goats.

There is one change in procedure that must be described. The subjects of the earlier experiment had used the conventional ESP symbols and had been required to make nine runs. These runs were divided into three sets, each set consisting of three runs, without intermission. Preliminary analysis indicated that at the beginning of the set of 75 responses the scores were above the level expected by chance, but that the scores fell to the chance level for the latter part of the 75 trials. The tentative inference was therefore drawn that the series was too long.

The 71 subjects tested in the first semester of the current experiment were required to follow this procedure, but later subjects made only 50 guesses as a unit instead of 75. In an effort to make the list of 50 items seem as short to the subjects as a list of 25, one further change was introduced: alternate items referred to colors (red, green, blue, yellow, and brown). Thus the list, from the subject's point of view, consisted of 25 items of the form "blue wave," "yellow star," etc. Hits were scored and counted on color and on symbol without making any distinction between them. Some classes completed four of these 50-item units, or 200 responses; others completed only three units within the class period. A few subjects

missed part of the session and thus made even fewer runs. In comparing the results, the mean ESP score was computed for each subject, and these means were used for the correlations reported in Table 5.

Picture-Frustration Study

For group administration of the Picture-Frustration Study, test booklets were distributed and the instructions were read aloud to the class. Completed papers were collected and were scored later by the experimenter. In all cases where the scoring was doubtful, a question about the response was written on the subject's paper. The papers were returned and the students were required to answer these questions. If the scoring still was unclear, individual inquiry was made. Scores were explained to the class, and the students were encouraged to check their own papers and to note any cases where they questioned or disagreed with the score. Papers were then collected for the second time and were rescored when necessary. All scoring of the picture test was completed without knowledge of the ESP results.

Ratings of Annoyance

Three methods were used to estimate how much frustration the subject experienced in making the ESP responses:

The first was an oral questionnaire. After each unit of 50 guesses, the subjects were asked to write at the bottom of their lists whether they felt relaxed, stimulated, or annoyed while making their guesses, or whether they reacted in some other way.

The second method, and the one which contributed most heavily to the final ratings, was a variant of the incomplete sentence technique described by Rotter, Rafferty, and Schachtitz (**5**). Instructions, at the top of a mimeographed sheet, were: "Complete these sentences to express your real feeling. Try to do every one. Be sure you write a complete sentence." Fourteen sentences were to be completed. They will be found in the sample records quoted at the end of this section. This test was administered toward the close of the experimental session, after the third unit of ESP responses.

A third indication of annoyance came from the subjects' answers to the question mentioned earlier: whether they felt that the success

of their ESP responses was determined by chance alone or that some other factor (which we call ESP) might be operative. As described in the previous report, the alternatives were put to the class at some length and were stated in a variety of ways. The subjects were encouraged to make a clear-cut choice between the two possibilities and then to write an explanatory paragraph describing the reason for the choice. The length of paragraphs and the vehemence with which ideas were expressed showed considerable variability.

The ratings of annoyance associated with the ESP test were completed without knowledge of the subjects' ESP scores. A seven-point scale was used. Ratings were based on the above three sources of information. The criteria were as follows:

Rating of 7: The subject expressed disapproval without a gleam of interest or enjoyment.

Rating of 6: The subject expressed annoyance with occasional sparks of interest, or with an admission of the possibility of interest or enjoyment.

Rating of 5: The subject expressed mild disapproval.

Rating of 4: Enjoyment and disapproval seemed equally balanced; or the subject seemed so uninterested that the experiment did not matter to him.

Rating of 3: Enjoyment or interest seemed slightly to outweigh disapproval or annoyance.

Rating of 2: The subject seemed clearly to enjoy the experiment, but made some qualifications or criticisms.

Rating of 1: The subject responded with great enjoyment and expressed no criticisms.

It was determined before the ratings were made that scores of 5 and 6, and only those scores, would be taken to indicate moderate annoyance with the ESP task. The subjects with these ratings, therefore, were the ones for whom it was expected to find a positive correlation between ESP scores and impunitive scores, and a negative correlation between ESP scores and extrapunitive scores.

In rating each subject on his level of annoyance with the ESP experiment, no attempt was made to force the scores into a predetermined distribution, nor was any attempt made to validate them.

In fact, by the nature of the method used, validation is impossible, for all factors that could have been used to check the ratings were instead used to determine them. Spot checking indicates that the ratings are highly reliable.

Sample records are given below for ratings of 7 through 4.

RECORD OF A SUBJECT WITH A RATING OF 7

Incomplete Sentences:

I feel . . . very irritable inside.
These experiments . . . seem to have no actual basis.
I want to know . . . just how such a thing is possible.
Next time . . . I'll draw better.
Telepathy . . . is scientifically unsound.
What annoys me . . . is the way this test is supposed to prove something.
A class period . . . was just wasted.
Probably clairvoyance . . . is just luck.
Sometimes . . . it works, but most of the time not.
I hate . . . to go to the country for the summer.
This school . . . would be great if not for some instigators.
Psychology really . . . is an excellent course.
I wonder . . . who's kissing her now.
Some teachers . . . have brilliant minds.

Attitude toward the possibility of paranormal success under the conditions of the experiment: I feel that if I got any right it was sheer accident. Telepathy is yet to be proven to me. Yet, somehow, I feel it might be so.

Response following each of the four units of the ESP guesses: Annoyed; annoyed; annoyed; indifferent.

(This subject was classed as a sheep because of the reservations expressed in "Yet, somehow, I feel it might be so." and in "Sometimes it [clairvoyance] works, but most of the time not." Individual inquiry confirmed that he did not reject *all* possibility of paranormal success. He made 29 correct ESP guesses in eight runs, which is 11 less than mean chance expectancy. Cases like this support the contention of Kahn (2) that attitude toward the possiblity of paranormal success is better measured by a scale with several levels of acceptance or rejection than by the two-level scale used in the present experiment.)

RECORD OF A SUBJECT WITH A RATING OF 6

Incomplete Sentences:

I feel . . . nervous, uneasy.
These experiments . . . are (would be) interesting; but they don't seem so today.
I want to know . . . what I want to know!!
Next time . . . let's go back to the text.
Telepathy . . . probably has something to it.
What annoys me . . . I don't really know.
(The rest of the record was left blank.)

Attitude toward the possibility of paranormal success under the conditions of the experiment: *Probably* something to it.

Response following each of the four units of ESP guesses: Nervous (on edge); same; annoyed; stimulated—this one interested me more.

(This subject was also classified as a sheep. His ESP successes totaled 66 in eight runs, or 26 above mean chance expectancy.)

RECORD OF A SUBJECT WITH A RATING OF 5

Incomplete Sentences:

I feel . . . good.
These experiments . . . are okay.
I want to know . . . why I am doing this.
Next time . . . I'll go away.
Telepathy . . . is okay.
What annoys me . . . most is this test.
A class period . . . is ending.
Probably clairvoyance . . . is not here.
Sometimes . . . I feel good.
I hate . . . someone.
This school . . . is okay.
Psychology really . . . teaches something.
I wonder . . . where I will go from here.
Some teachers . . . are okay.

Attitude toward the possibility of paranormal success under the conditions of the experiment: I am definitely open to persuasion, to be shown that telepathy is possible. There is some belief for it in my mind. Sheep.

Response following each of the four units of ESP guesses: My mind was tense. I was straining it; still straining; still tense; tense.

(The subject was a sheep. His ESP score was 44 for eight runs.)

RECORD OF A SUBJECT WITH A RATING OF 4

Incomplete Sentences:

I feel . . . tired.
These experiments . . . are interesting.
I want to know . . . what my IQ is.
Next time . . . I'll finish my work sooner.
Telepathy . . . in my opinion can exist.
What annoys me . . . is someone who is late.
A class period . . . in psych is too short.
Probably clairvoyance . . . is not possible.
Sometimes . . . I'd like just to curl up and go to sleep.
I hate . . . costume jewelry.
This school . . . is wonderful.
I wonder . . . if I'll pass all my courses.
Some teachers . . . are horrible, but others are wonderful.

Attitude toward the possibility of paranormal success under the conditions of the experiment: Impossible.

Responses following each of the four units of ESP guesses: relaxed; relaxed; indifferent; indifferent.

(Here the attitude expressed seemed to be a tolerant unconcern; the experiments were interesting but "impossible." The subject was, of course, classified as a goat, because she did not accept the possibility of paranormal success *under the conditions of the experiment,* even though she believed that telepathy was possible. Her ESP score was 41 for eight runs.)

RESULTS

Of the 272 subjects who took part in the ESP experiments, five did not fill out the Picture-Frustration Study, and one began but refused to complete it. Only the data of the remaining 266 subjects will be cited here.

ESP Scores

In previous experiments by the author, sheep have made higher average ESP scores than goats. The same tendency appears here. When the total runs made by the sheep are pooled and compared with the corresponding total for the goats, the total deviation is +200 for the sheep group and −61 for the goat group (Table 1).

Table 1

GENERAL RESULTS OF SHEEP AND GOATS

	Number of Subjects	Number of Runs	Hits	Dev.	CR	P
Sheep	162	1250	6450	+200	2.83	.002
Goats	104	831	4094	− 61	1.06	> .1

Table 2

MEAN SCORES OF SHEEP COMPARED WITH MEAN SCORES OF GOATS

	Number of Subjects	Mean of Subjects' Mean Scores	t_d	P
Sheep	162	5.15	2.37	.01
Goats	104	4.93		

The difference between the two groups was evaluated by calculating the mean score of each subject and testing the significance of the difference between the distributions of mean scores. With this method, the difference between the groups is significant at the level of $P = .01$ (Table 2).

It should also be noted that 54% of the sheep (87 subjects) and 46% of the goats (48 subjects) had more hits than mean chance expectancy. The difference between these percentages is not significant.

Since the psychological dynamics of making ESP responses may well have been different for the two groups, their data have been presented separately in later tables.

Table 3

SCORES ON THE PICTURE-FRUSTRATION STUDY (IN PERCENTAGES)

	ROSENZWEIG'S NORMS		SCORES IN THIS EXPERIMENT		
	Males (n=236)	Females (n=224)	Sheep (n=162)	Goats (n=104)	Total (n=266)
Extrapunitive	45%	45%	45%	46%	46%
Intropunitive	28	28	28	27	28
Impunitive	27	28	27	27	27

Picture-Frustration Study

Picture-Frustration scores obtained in this experiment are summarized in Table 3. Norms reported by Rosenzweig (**4**) are listed for comparison. It is obvious that the means of our subjects are approximately the same as those for the standard group he describes. Differences between sheep and goats in respect to Picture-Frustration scores are negligible.

Ratings of Annoyance

The distribution of annoyance ratings is given in Table 4. As was to be expected, the patterns are different for the sheep, who intellectually accepted the possibility of success in their responses, and for the goats, who rejected this possibility. There is considerable overlap between the two groups. Perhaps the overlap indicates the degree to which non-intellectual factors determined their total attitude and affect.

Table 4

DISTRIBUTION OF RATINGS OF ANNOYANCE

Rating	Number of Cases	
	Sheep	Goats
7 (high)	26	42
6	36	30
5	38	14
4	36	12
3	14	5
2	11	1
1	1	0

Correlations between ESP Scores and Picture-Frustration Scores

Table 5 gives the data from all subjects who have previously taken these tests, from all of the current subjects, and from the group with which we are most concerned—namely, those who were annoyed, but not intensely annoyed, by the experiment and who were accordingly given annoyance ratings of 5 or 6 on a 7-point scale.

The data of the current group, taken as a whole, show a very slight tendency to score in the same direction as previously tested subjects. This effect is not significant. But when the group with annoyance ratings of 5 or 6 is considered, the correlations are

Table 5

CORRELATIONS BETWEEN ESP SCORES AND RATINGS ON THE PICTURE-FRUSTRATION STUDY

Subjects	Number of Subjects	Correlation of ESP Score and Extrapunitive Rating			Correlation of ESP Score and Intropunitive Rating			Correlation of ESP Score and Impunitive Rating		
		r	t	P	r	t	P	r	t	P
All Earlier Subjects										
Sheep	345	−.14	2.56	.005	+.04	.74	> .1	+.15	2.84	.002
Gaots	150	−.06	.74	> .1	+.15	1.83	.03	+.04	.51	> .1
Total	495	−.12	2.60	.005	+.06	1.33	.09	+.12	2.71	.003
All Current Subjects										
Sheep	162	−.03	.38	> .1	−.01	.13	> .1	+.06	.76	> .1
Goats	104	−.11	1.12	> .1	+.07	.71	> .1	+.05	.50	> .1
Total	266	−.07	1.14	> .1	+.04	.65	> .1	+.05	.81	> .1
Current Subjects Who Were Moderately Frustrated										
Sheep	74	−.27	2.39	.01	+.04	.33	> .1	+.30	2.68	.004
Goats	44	−.21	1.40	.08	+.25	1.67	.05	+.12	.79	> .1
Total	118	−.23	2.55	.005	+.12	1.29	.1	+.21	2.31	.01

unmistakably in the predicted direction and are significantly different from chance for the relation between ESP scores and extrapunitive ratings ($t = 2.55$; $P = .005$) as well as for the relation between ESP scores and impunitive ratings ($t = 2.31$; $p = .01$).

DISCUSSION

The crucial test for our hypothesis about the relation between ESP scores and scores on the Picture-Frustration Study consisted of examining the records of the moderately frustrated subjects. The prediction was made, on the basis of previous findings, that those whose habitual response to mild frustration was extrapunitive would be aggressive and hostile while making the ESP responses and would therefore make low ESP scores. Conversely, it was predicted that those whose characteristic response to mild frustration was impunitive would emphasize the interesting or enjoyable aspects of the experiment, would react more favorably and coopera-

tively, and would therefore make higher ESP scores. These predictions were supported by the data of Table 4.

The other part of our hypothesis was that, because of the nature of the Picture-Frustration Study, Picture-Frustration scores would show a meaningful relation only for the moderately frustrated group. Comparison of the low non-significant correlations for "All Current Subjects" in Table 5 with the higher, significant correlations for "Current Subjects Who Were Moderately Frustrated" tends to confirm this prediction also.

It will be noted that the correlations between ESP score and intropunitive (that is, self-punishing) ratings have been calculated. Some of them are suggestive, although none reaches the .01 level of significance. It is of some interest that they indicate a positive relationship between intropunitive scores and ESP success for the goats, that is, for the subjects who condemn the ESP experiment as attempting to study something which could not occur. Let us consider what this might mean.

A high intropunitive rating implies that in a moderately frustrating situation a subject tends to blame himself for the frustration, or to feel embarrassed about it, or to take upon himself the responsibility for making things go right. He is likely, then, not to be cocksure. When moderately frustrated because he disagrees with the experimenter's point of view, he may tend to deprecate his own opinions or to make a special effort to cooperate with the experimenter. We expect most goats to hope for failure; but perhaps *intropunitive* goats, if moderately frustrated, will try particularly hard for success just because they do not want it themselves. It seems worth while, therefore, to state as a tentative hypothesis that goats who are moderately frustrated by an ESP experiment and who tend, when moderately frustrated, to disparage their own opinions or to take the responsibility for trying to set things right, will make higher ESP scores than goats who do not tend, under moderate frustration, to disparage their own opinions or to take such responsibility.

The reader may have wondered, in looking over the distribution of annoyance ratings in Table 4, that so much annoyance was expressed or implied. The author has also wondered at it and suggests two possible reasons for this unexpected finding. One is

that the use of the sentence completion technique may have influenced the results. This projective test taps deep underlying attitudes in addition to the mood of the moment. The resentment which was expressed may relate in part to feelings about home relationship or college authority patterns, as well as to feelings about the experimental situation. Another reason, almost surely, lies in the rigid controls that are routine in ESP research: the concealed targets, the repetitive guessing at symbols that are convenient for statistical treatment but are dull from the subject's point of view, the response sheets that must be collected before the targets can be shown to the subjects, and so on.

In individual ESP tests it is, of course, possible to create such a friendly or challenging atmosphere that the whole situation becomes agreeable, for most subjects. In group testing, however, it may seem necessary to choose between having so few responses from each subject that the data are sparse, or else having the session so long that it tends to be dull. It seems probable that the subjects' boredom or annoyance in this experiment caused an emotional withdrawal from the prolonged ESP test, and that the emotional withdrawal, in turn, resulted in ESP scores that were closer to chance expectancy than if responses had been made with a high level of interest, and with consistent emotional involvement. On the basis of this reasoning we can formulate another hypothesis: that whatever trends appear in research like the present one, where the subjects tend to be bored, will appear more clearly in an experiment where the subjects' interest is maintained at a consistent level. Different trends, of course, are to be expected, depending on whether the consistent level of interest is high or low.

References

1. Eilbert, L., and Schmeidler, G. R. A study of certain psychological factors in relation to ESP performance. *J. Parapsychol.*, 1950, **14**, 53-74.
2. Kahn, S. D. Studies in extrasensory perception. Experiments utilizing an electronic scoring device. *Proc. Amer. Soc. psych. Res.*, 1952, **25**, 1-48.
3. Rosenzweig, S. The picture-association method and its application in a study of reactions to frustration. *J. Personal.*, 1945, **14**, 3-23.
4. ———. Revised norms for the adult form of the Rosenzweig Picture-Frustration Study. *J. Personal.*, 1950, **18**, 344-46.

5. ROTTER, J. B., RAFFERTY, J. E., AND SCHACHTITZ, E. Validation of the Rotter Incomplete Sentences. *J. consult. Psychol.*, 1949, **13**, 348-56.

6. SCHMEIDLER, G. R. Personality correlates of ESP as shown by Rorschach studies. *J. Parapsychol.*, 1949, **13**, 23-30.

7. ———. Some relations between Picture-Frustration ratings and ESP scores. *J. Personal.*, 1950, **18**, 331-43.

8. SCHMEIDLER, G. R., AND MURPHY, G. The influence of belief and disbelief in ESP upon ESP scoring level. *J. exp. Psychol.*, 1946, **36**, 271-76.

City College of New York
New York, N. Y.

CLOCK CARDS AND DIFFERENTIAL SCORING TECHNIQUES

By A. M. J. Mitchell and G. W. Fisk

ABSTRACT: This report summarizes several experiments in which ESP and PK tests were evaluated by two methods. The first was the one commonly used, in which a trial is scored either right or wrong. The second was a differential scoring technique in which calls were graded from a direct hit, through a near-miss, down to a complete failure. For the ESP tests, the target cards were made to resemble a clock face showing an hour hand pointing to one of the 12 hours, but with no minute hand. In the differential method a hit on the hour selected as target received a value of 0; the two adjacent hours received 1 point against them; and so on, until the hour opposite the target received 6.

A group of 16 subjects, on 2832 trials calling down through decks of 12 Clock cards in sealed envelopes, did not score significantly by the usual method of counting direct hits only, but they did score significantly by the differential method. In a similar series comparing the results of two different experimenters, one experimenter had statistically significant results by both scoring methods, but the difference between the two experimenters was not significant.

For the differential method of evaluating dice tests, the target face was valued at 0; the adjacent faces at 1; and the opposite face at 2. In 39,992 die-throws, the 10 subjects gave significant results by the customary method and even higher results on the differential method. A variation of this procedure, in which each of the six die-faces was given a separate value, did not give significant results.

The tests were exploratory insofar as the comparison of the two scoring methods is concerned. While there was a general tendency for the differential method to yield more significant results than scoring direct targets alone, the data are not adequate for a definite conclusion on this point.

Mr. Mitchell is a physicist employed on telecommunications work; Mr. Fisk is retired from business; both are members of the London Society for Psychical Research and have for the last few years been carrying out the experimentation summarized in the following article.—Ed.

Introduction

During 1953 three papers describing ESP and PK tests appraised by new scoring techniques were published in the *Journal of the Society for Psychical Research* (**1**, **2**, **4**). The purpose of this article is to outline the methods used, summarize the results obtained, and indicate areas requiring further research. Full experimental and mathematical details will not be repeated here since anyone interested can consult the original papers; and the space at our disposal can therefore be devoted, we hope more profitably, to the discussion and

interpretation of the results obtained. The general differential scoring technique will be discussed first, as it is a common feature of the ESP and PK experiments described later in the paper.

Differential Scoring

In the usual ESP tests the outcome of a trial is either a hit or a miss and is scored 1 or 0. Now, it seems reasonable to suppose that psi processes are not all-or-nothing phenomena and that some of those trials which do not result in a hit may be near-misses. An analogy which may be found helpful is that of an observer seeking to assess the ability of a marksman on a rifle range. In firing at the target the marksman may sometimes score a bullseye, that is, a complete success; but "inners," "magpies," and "outers" are also taken into account in varying degree when the result of the shoot is assessed. This scoring method is clearly a more sensitive measure of marksmanship than one in which only the bullseyes are counted, and it may provide clues as to the shooting characteristics of the marksman, the bias of the rifle, prevailing weather conditions, etc.

To devise a differential scoring technique in ESP three main factors must be considered. First, one must define the number of categories into which the outcome of a trial may fall. One such category will clearly be the direct hit where the call agrees with the target; it must then be decided what will be recognized as an "inner," "outer," etc. Secondly, one must choose appropriate scores to be given to each category. In our work we adopted the convention of scoring the "error" of a trial. For example, a direct hit or "bullseye" is scored 0; an "inner" is 1; a "magpie" is 2; an "outer" is 3. Finally, the statistics of the score distribution must be evaluated, and here it is advantageous to set up a scoring technique for which there are convenient tables available for computing the chance probability. The method of calculating the shape of the score distribution is detailed in Appendices to references (**1**) and (**4**). The potential sensitivity of differential scoring systems over the usual direct-hit method is also illustrated. The theoretical distributions for scores under the systems mentioned below are sufficiently close approximations to the normal curve to permit use of the usual CR probability tables in evaluating the results. It should be noted that the division into categories and the scores to be given are decided

arbitrarily and form a hypothesis to be tested by application to experimental data, as indeed does the more usual scoring technique.

Clock Cards

Test Material and Scoring System

The test material comprises a deck of cards, each bearing the diagram of a clock face with an hour hand pointing exactly to one of the hours I-XII. The minute hand is omitted (see Fig. 1).

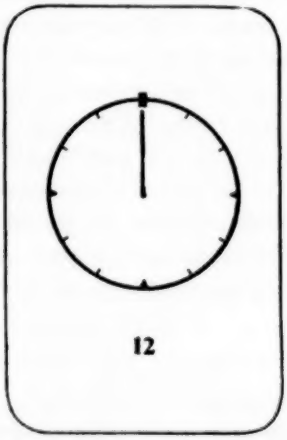

Fig. 1. Sample of Clock cards used as targets.

With the cards arranged in random order the subject attempts to guess the hour shown by each card in turn and so has the choice of twelve hours in making his call. Suppose, for example, the target card depicts 12 o'clock. Using the usual scoring method a subject calling XII would be credited with a hit, score 1; but if he called any other hour his score would be 0. Using the differential technique a subject calling XII is again credited with a hit but, on our convention, since he has made no error, the score will be 0. Now, if he called I or XI his divergence (Δ) from success would be plus or minus one hour and in either case would be scored 1. It may be regarded as an angular error of 30°. If he called II or X, the divergence would be plus or minus two hours and scored 2. The angular error is 60°. Similarly the scores for the remaining possible calls are 3, 4, 5, 6, this last, corresponding to a call of VI, represents a complete failure, an angular error of 180°.

The probabilities of calling the above divergences are 1/12 for Δ_0, 1/6 each for Δ_1 to Δ_5, and 1/12 for Δ_6. The chance expectation of the sum of the scores (divergences) in n trials is $3n$ with a standard deviation of $\sqrt{(19n/6)}$. It will be appreciated that a divergence sum smaller than $3n$ denotes a positive deviation, that is, a score above chance expectation.

Preliminary Tests

The possibilities of the Clock technique were first explored by G.W.F., who used a watch with one hand removed as the target (1). After setting the hand to a randomly selected hour he exhibited the back of the watch to two young subjects who guessed the time, and he recorded their calls. Although the experimental conditions were not rigid, the results were most encouraging and are summarized in Table 1. In Table 1 a comparison is made between the usual direct-hit scoring system and the differential scoring system outlined above. In order to conform with normal statistical practice the symbol t is used throughout this paper to denote the critical ratio.

Table 1

SUMMARY OF GESP TRIALS

	Direct Hits	Divergence Hits						Sum of Divergences
	Δ_0	Δ_1	Δ_2	Δ_3	Δ_4	Δ_5	Δ_6	Δ_1 - Δ_6
Subject S. B. (1440 Trials)								
No. of Hits	161	269	224	246	221	222	99	4,043
M. C. E.	120	240	240	240	240	240	120	4,320
Dev.	+41	+29	−16	+6	−19	−18	−21	+277
t	3.91	2.05	1.13	0.43	1.34	1.27	2.005	4.10
P	9.2×10^{-5} (approx.)	0.04	0.26	....	0.18	0.20	0.045	4.0×10^{-5} (approx.)
Odds	10,845:1 (approx.)	....	....	....	....	....	....	24,000:1 (approx.)
Subject J. B. (1524 Trials)								
No. of Hits	125	280	259	247	238	252	123	4,489
M. C. E.	127	254	254	254	254	254	127	4,572
Dev.	−2	+26	+5	−7	−16	−2	−4	+83
t	...	1.79	....	....	1.10	....	....	1.195
P	...	0.073	....	....	0.27	....	....	0.23

It will be seen that subject S.B. attained a significant score by both systems primarily because of a large number of direct hits. However she also had a small excess of near-misses (Δ_1) and a deficit of calls which were four or more hours from the target. The differential scoring system which was used takes account of this and so leads to a more highly significant result than that given by the direct-hit method. J.E. owed her small positive score chiefly to an excess of near-misses, but the effect is insufficiently marked to produce statistical significance. It may be mentioned that both girls had been tested previously on several occasions with ordinary ESP cards, but without success.

"Down Through Sealed Decks" (DTSD) Trials

Encouraged by these results G.W.F. arranged more refined tests using printed Clock cards (**1**). Sealed decks of 12 Clock cards, arranged in random order by means of random numbers (open decks), were distributed in sets of four runs at a time (48 trials) from May, 1951, to October, 1952, to 54 subjects. Most of these subjects had already taken part in tests ordinary ESP cards and none had shown any consistent scoring above chance. They were instructed to call the hour depicted on each card from top to bottom of the deck ("Down Through Sealed Decks" or DTSD). Score sheets and the still unopened decks were returned to G.W.F. for checking.

The grouped results showed a highly significant positive deviation. The probabilities were: for the direct hit score, $P = 0.006$; and for the divergence score, $P = 1.6 \times 10^{-6}$.

It so happens that some of the 54 subjects were enlisted by others to whom the decks had been originally sent so that several subjects called the same decks. There are statistical objections to grouping results in this event, and so we quote here the results for those primary subjects (17 in number) to whom separate decks were sent. The results are given in Table 2, which is subdivided to show separately the results obtained by one outstanding subject.

The group of 16 subjects scored a small excess of direct hits, but not enough to reach significance. They did, however, gain a large surplus of near-misses which was sufficient to raise their dif-

Table 2

SUMMARY OF NON-STACKING DTSD RESULTS

(16 Subjects, 2832 Trials)

	Direct Hits	Divergence Hits						Sum of Divergences
	Δ_0	Δ_1	Δ_2	Δ_3	Δ_4	Δ_5	Δ_6	$\Delta_1 - \Delta_6$
No. of Hits	247	540	466	479	438	451	211	8,182
M. C. E.	236	472	472	472	472	472	236	8,496
Dev.	+11	+68	−6	+7	−34	−21	−25	+314
t	0.747	3.43	0.30	0.35	1.71	1.06	1.70	3.32
P	0.45	0.0006	....	....	0.09	0.29	0.09	0.0009
Odds	...	....	....	....	....	....	....	1,100:1
	(Chi-square = 16.521; P = 0.011 approx.)							

SUBJECT S. M. (240 Trials)

No. of Hits	40	49	47	34	33	26	11	573
M. C. E.	20	40	40	40	40	40	20	720
Dev.	+20	+9	+7	−6	−7	−14	−9	+147
t	4.67	1.56	1.21	1.04	1.21	2.43	2.10	5.33
P	3.4×10^{-6} (approx.)	0.12	0.23	0.30	0.23	0.015	0.036	1.0×10^{-7} (approx.)
Odds	294,000:1 (approx.)	....	....	....	....	....	....	10 million to 1 (approx.)
	(Chi-square = 34.325; P = 5.1×10^{-6} approx.)							

ferential score to a significant level. This is the effect, noted on a smaller scale, in the data of the subject J.B. above.

S.M.'s score is outstanding because of her excess of direct hits, but here again, because this is coupled with a deficit of calls which were three or more hours from the target, an even higher significance level is reached using the differential scoring method.

A Dual Experiment with Clock Cards (4)

In discussing some puzzling effects in the Fisk data obtained in some previous group ESP experiments (3), Dr. D. J. West suggested that G.W.F. seemed to belong to a group of successful experimenters while he himself belonged to a group who could only achieve null results. In view of the apparent greater sensitivity of the Clock card technique compared with ordinary ESP cards, it was decided to run a test with G.W.F. and D.J.W. acting alternately

as experimenters with the same subjects. The DTSD method described above was followed. Thirty-two sealed decks of twelve cards each were sent in sets of four decks at a time to each of 20 subjects. Sixteen of these decks were made up by G.W.F. and 16 by D.J.W. The order in which the subjects were given the two kinds of decks was randomized, and precautions were taken to prevent subjects knowing (by normal means at least) which experimenter had made up any particular deck. The results of this test are given in Tables 3 and 4.

Table 3

DIRECT HIT SCORES

(Each subject made 192 trials for each experimenter)

Subject	FISK TARGETS			WEST TARGETS			COMBINED		
	Dev.	*t*	P*	Dev.	*t*	P	Dev.	*t*	P
J. C.	−3	−0.78	.....	−6	−1.57	.....	−9	−1.66	
J. B.	13	3.39	0.0007	−6	−1.57	.....	7	1.29	
S. B.	10	2.61	0.009	−1	−0.26	.....	9	1.66	
Dr. B.	12	3.13	0.002	2	0.52	.....	14	2.59	0.01
P. C.	1	0.26	.....	0	0	.....	1	0.18	
E. D.	4	1.04	.....	−5	−1.31	.....	−1	−0.18	
C. R.	10	2.61	0.009	4	1.04	.....	14	2.59	0.01
I. F.	5	1.31	.....	0	0	.....	5	0.92	
E. W.	0	0	.....	0	0	.....	0	0	
I. J.	7	1.83	.....	3	0.78	.....	10	1.85	
C. L.	−1	−0.26	.....	1	0.26	.....	0	0	
B. O.	6	1.57	.....	0	0	.....	6	1.11	
G. R.	−5	−1.31	.....	−1	−0.26	.....	−6	−1.11	
A. W.	2	0.52	.....	3	0.78	.....	5	0.92	
S. M.	8	2.09	0.04	20	5.22	10^{-7}	28	5.17	10^{-7}
J. W.	−6	−1.57	.....	1	0.26	.....	−5	−0.92	
S. G.	1	0.26	.....	−4	−1.04	.....	−3	−0.55	
D. J.	−1	−0.26	.....	0	0	.....	−1	−0.18	
H. G.	0	0	.....	4	1.04	.....	4	0.74	
K. G.	2	0.52	.....	−1	−0.26	.....	1	0.18	
Total	65	3.79	0.00015	14	0.82	.....	79	3.26	0.0011
Total Without S. M.	57	3.41	0.0065	−6	−0.36	.....	51	2.14	0.032

*The probabilities associated with individual subjects are only approximate because, with relatively small samples and rather large *t* values, the normal distribution is not strictly applicable.

The main purpose of these tests was to see whether subjects reacted differently to decks of cards made up, and afterwards assessed, by the two experimenters. Although the results at first sight suggest that G.W.F. gets the highest scores from this group of subjects, a statistical analysis of the difference between a subject's score on G.W.F.'s decks and those of D.J.W. does not confirm

Table 4

DIVERGENCE SCORES

Subject	FISK TARGETS			WEST TARGETS			COMBINED		
	Dev.	*t*	P*	Dev.	*t*	P	Dev.	*t*	P
J. C.	−28	−1.13	.....	−44	−1.78	.....	−72	−2.06	0.039
J. B.	81	3.28	0.001	−54	−2.19	0.03	27	0.77	
S. B.	33	1.34	.....	−12	−0.49	.....	21	0.60	
Dr. B.	84	3.41	0.0006	32	1.30	.....	116	3.33	0.0009
P. C.	40	1.62	.....	−19	−0.77	.....	21	0.60	
E. D.	35	1.42	.....	−23	−0.93	.....	12	0.34	
C. R.	56	2.27	0.02	− 8	−0.32	.....	48	1.38	
I. F.	−33	−1.34	.....	5	0.20	.....	−28	−0.80	
E. W.	−17	−0.69	.....	−27	−1.10	.....	−44	−1.26	
I. J.	16	0.65	.....	2	0.08	.....	18	0.52	
C. L.	6	0.24	.....	6	0.24	.....	12	0.34	
B. O.	57	2.31	0.02	18	0.73	.....	75	2.15	0.03
G. R.	−24	−0.97	.....	− 6	−0.24	.....	−30	−0.86	
A. W.	−19	−0.77	.....	−25	−1.01	.....	−44	−1.26	
S. M.	110	4.46	10^{-5}	117	4.75	10^{-6}	227	6.51	10^{-10}
J. W.	7	0.28	.....	6	0.24	.....	13	0.37	
S. G.	27	1.09	.....	−12	−0.49	.....	15	0.43	
D. J.	−22	−0.89	.....	11	0.44	.....	−11	−0.32	
H. G.	−25	−1.01	.....	16	0.65	.....	− 9	−0.26	
K. G.	− 7	−0.28	.....	−44	−1.78	.....	−51	−1.46	
Total	377	3.42	0.0006	−61	−0.55	.....	316	2.02	0.04
Total Without S. M.	267	2.48	0.013	−178	−1.66	.....	89	0.57	

*The probabilities associated with individual subjects are only approximate because, with relatively small samples and rather large *t* values, the normal distribution is not strictly applicable.

this view. The difference is negligible taking into account the direct hits only, but is approaching significance ($P = 0.03$) on the differential scoring system.

There were, however, some interesting and significant scores by individual subjects. S.M., who had previously done well with Clock cards (see Section on DTSD trials above), again produced a striking score with both experimenters. In this case D.J.W.'s decks were actually the rather more favored, and her scores were more significant when assessed on the differential system than on the usual direct-hit method customarily used. Dr. B., on the other hand, preferred G.W.F.'s decks and also obtained more significant results by the differential method. J.B. and S.B., the two young guinea-pigs in the preliminary Clock tests, obtained significant results, but now it is J.B. who is the higher scorer. They did better with G.W.F.'s decks but obtained more highly significant results when appraised by the direct-hit method. It should be remembered,

however, that the preliminary tests with S.B. and J.B. followed a GESP technique, whereas these later ones were DTSD. Subjects often show very different responses to the two methods. Dr. Soal's subject, Mrs. Gloria Stewart, for example, who was so outstanding in GESP tests failed completely in DT tests.

General Conclusions from the Clock Card Results

It must be said immediately that only tentative conclusions concerning the relative advantages of the two methods of scoring can be drawn from this set of data. Certain subjects have obtained significant scores with Clock cards, some of them in two rather different tests. In some subjects it does appear that the differential scoring system is a more sensitive measure of appraising their ESP ability. This means, in effect, that they do tend to guess near, as well as on, the target hour and avoid calling an hour diametrically opposite to the target. This effect may be similar to that which would occur in the visual perception of the hour hand under conditions of poor visibility or by viewers suffering from a pronounced astigmatism. This view could be tested experimentally, as was suggested to us by J. Fraser Nicol in a private communication. Although the differential scoring system used is successful in some instances, the observed dispersion of divergences suggests that a more efficient valuation might be possible. Perhaps the outcome of a call has been divided into too many categories. For example, the relative unimportance of Δ_2 and Δ_3 is noticeable with many subjects. Possibly the substitution of the clock face by a "compass face" with eight alternatives (N, NE, E, SE, S, SW, W, and NW) in lieu of the twelve hours might be worth trying, although many subjects would not be so familiar with the points of the compass as with the hours of the clock. Perhaps the most important advantage of differential scoring over the usual hit-or-miss method will lie in this ability to distinguish between various ways in which subjects achieve a significant score (either positive or negative). For example, the excess or deficit may be due to the number of direct hits only, or it may be entirely due to some particular value of a near-miss or relative failure. When setting up a more efficient scoring system the possibility must not be overlooked that the system may have to be tailored to suit individual percipients.

One final point with regard to Clock cards must be mentioned, although it does not bear directly on the results described above. It is the fact that subjects do not record each of the target hours with equal frequency. There are very strong deviations from chance expectation. Grouped results to date show that the hours XII, XI, X, VIII (in order of dislike) are called less frequently, while I, IX, II, III (in order of preference) are favorites.

ESP/PK Tests with Differential Scoring

Following the success of the early Clock card tests, A.M.J.M. suggested that differential scoring might be applied to PK tests with dice. When a die is thrown for a target, the die can come to rest in one of three main orientations. The die face agreeing with the target may be found uppermost—a direct hit; or in a vertical position; or in contact with the surface upon which the die is thrown. If we have a subject throwing a die for a target, these three outcomes of a single trial could be regarded as representing successively greater orientational errors. In the Die Orientation (DO) scoring method these outcomes are scored 0, 1, and 2 respectively. They occur with nominal probabilities of 1/6, 4/6, and 1/6. In n throws the mean chance expectation of the total score is n with a standard deviation of $\sqrt{(n/3)}$.

This scoring method was tested in experiments arranged by G.W.F. (2). The basic experiment comprised the throwing of a number of dice for a succession of targets, each displayed for 24 hours and changed at approximately 8 A.M. daily by G.W.F. in Surrey, England. The subjects were situated between three and 300 miles from the targets, which were unknown to them. Under these conditions we presumed that significant scores would be due to a combination of ESP and PK. Over the majority of the tests G.W.F. selected the targets, using a table of random numbers. Following outstanding results with one subject, Dr. B., who obtained significant results with Clock cards (see Section on Dual Experiment above), a further similar test was run; but this time Dr. West selected the targets, using a 6 x 6 Latin Square, and G.W.F. was kept in ignorance of the targets until the scores were assessed. The results of these tests are shown in Table 5.

The grouped results are significant and more highly so when the

Table 5

SUMMARY OF RESULTS FOR TEN SUBJECTS ON ESP/PK TESTS

Subject	Location	Target dist. miles	Total throws n	Freq. of Occurrence (Dev. from MCE) for Scores 0, 1, 2			Scoring Method: Direct Hits		Scoring Method: Die Orientation	
				0	1	2	t	P	t	P
S. M.	Birmingham	110	600	+ 7	+19	− 26	+0.77	0.441	+2.33	0.020
P. V.	France	160	646	+ 20.3	− 1.6	− 18.6	+2.14	0.032	+2.66	0.008
C. V.	France	160	646	+ 9.3	−21.6	+ 12.3	+0.98	0.327	−0.20	0.841
Dr. B. (1)	Devonshire	170	10,000	+117.3	−13.6	−103.6	+3.15	17×10^{-4}	+3.83	1.3×10^{-4}
A. N.	London	15	2,700	− 14	+10	+ 4	−0.72	0.471	−0.60	0.549
A. I. B.	Devonshire	170	900	+ 2	+15	− 17	+0.18	0.857	+1.10	0.271
B. R. B.	Devonshire	170	900	− 27	+24	+ 3	−2.41	0.016	−1.73	0.084
J. A. C.	Kingston	3	10,000	+ 0.3	+30.3	− 30.6	+0.01	0.99	+0.54	0.589
Dr. B. (2)*	Devonshire	170	3,600	+ 53	−52	− 1	+2.32	0.020	+1.56	0.119
P. A. C.	Liverpool	200	5,000	+ 49.6	−16.3	− 33.3	+1.96	0.050	+2.03	0.042
J. W.	Eire	320	5,000	+ 54.6	−41.3	− 13.3	+2.07	0.038	+1.66	0.097
Total			39,992	+272.6	−48.3	−224.3	+3.66	2.6×10^{-4}	+4.31	0.2×10^{-4}

*Latin Square targets.

Chi Square (2 d.f.) = 18.79
Odds greater than 10,000:1 against

Odds 4,000:1 against

Odds 50,000:1 against

data are appraised by the Die Orientation differential method. Inspection of the frequency-of-occurrence column shows that this is due to an excess of direct hits together with a deficit of occasions when the target face of the die came to rest in contact with the throwing surface. This effect does not occur with every subject but is particularly marked in the first series with Dr. B. as subject. In her second series, however, although a high score was obtained, it was almost entirely due to an excess of direct hits. Whether this was due to the use of Latin Square targets, the collaboration of Dr. West, or some other reason, is a matter of speculation. Dr. B. was then in failing health and shortly afterwards had a very serious breakdown.

A further group of tests were devised and arranged by A.M.J.M. Briefly, subjects were asked to throw dice alternately for sixes and ones. They were told that when throwing for sixes, for example, most weight would be given when they obtained a six. Some weight would be given for fives and, successively, less weight for lower faces. When throwing for ones, most weight would be given for a one, less weight for a two, and so on. A differential scoring system was worked out on this basis and applied to the resulting data. The results were quite insignificant. The subjects used for these tests

had taken part in ESP tests without success and none of them had ever shown any ability in PK trials. One can only say they ran true to form.

Conclusions

A summary has been made of some ESP and PK tests which were designed to be assessed by two methods. The first was the one normally adopted, in which the results of a trial is scored as a hit or a miss. The second employed a differential scoring technique in which the calls were graded from a direct hit, through the near-miss, down to a complete failure.

Special cards depicting a clock face with one hour hand were used for the ESP tests and normal dice were employed for the PK tests. A common feature of all the tests was that in general the subjects were separated some considerable distances from the experimenter, and from each other. In the PK tests the subjects were also some distance from the targets.

Although no far-reaching conclusions can be drawn from these exploratory tests a number of outstanding scores were obtained by several subjects, and in some cases the significance is appreciably greater when the results are assessed by a differential scoring method than by the usual method of counting direct hits only.

It will be noted that this work has been carried out almost entirely by a small group of two or three experimenters. It is hoped that the results obtained may be considered sufficiently encouraging to stimulate others to explore differential scoring techniques and throw some light on the peculiar effects revealed.

References

1. Fisk, G. W., and Mitchell, A. M. J. ESP experiments with Clock cards. A new technique with differential scoring. *J. Soc.psych. Res.*, Lond., 1953, **37**, 1-14.
2. Mitchell, A. M. J., and Fisk, G. W. The application of differential scoring to PK tests. *J. Soc. psych. Res.*, Lond., 1953, **37**, 45-61.
3. West, D. J. Home-testing ESP experiments. An examination of displacement effects. *J. Soc. psych. Res.*, Lond., 1953, **37**, 14-25.
4. West, D. J., and Fisk, G.W. A dual ESP experiment with Clock cards. *J. Soc. psych. Res.*, Lond., 1953, **37**, 185-97.

2 Ditton Grange Close
Ditton Hill
Surrey, England

A REVIEW OF THE PEARCE-PRATT DISTANCE SERIES OF ESP TESTS

By J. B. Rhine and J. G. Pratt

A number of considerations have contributed to our decision to present the original and subsequent work identified with what has come to be known as the Pearce-Pratt Distance Series of ESP tests, carried out in 1933-34 at the Parapsychology Laboratory at Duke University. One reason for the review is the need expressed by some students of the subject for a more complete and detailed account of the orignal experiment than is to be found in any one publication. The first part of the series, what is known as Subseries A, was published in the monograph *Extrasensory Perception* written in 1934 by J. B. R. (**6**). This section was all that was completed at the time the monograph was written. In 1936 a brief account of the series and its total results was given in an article by J. B. R. in the *Journal of Abnormal and Social Psychology* (**8**), and in 1937 a condensed version of this article was included in the first number of the *Journal of Parapsychology* (**7**).

Another reason for the present undertaking is the fact that almost immediately upon publication the Pearce-Pratt Series received special attention. It represented a methodological advance over earlier experimental work in parapsychology; and both for the laboratory group associated with the experiment and for those who were attempting to appraise and criticize the evidence for extrasensory perception, the series had to be considered. Morover, as new questions were raised about the series, further analyses of the data resulted. Most of these analyses were reported as they were completed, but to the student of today it would be a difficult undertaking to run them all down.

There is the further point that it is now possible to appraise the experiment and its results in the light of the developments of the intervening twenty years, the most productive period of parapsychology. It was considered an advantage to older students as well as new, therefore, for the authors to assemble for re-examination the

factual matter that has accumulated around this single experimental series.

Something should be said regarding the general background of the research. First, there is the all-important aspect of personnel. It should not be forgotten that without Prof. William McDougall's appreciation of the problem and his tolerant and courageous interest in seeing it investigated under good conditions in a psychology laboratory, the experiment would not have been possible. J.B.R. was at the time an assistant professor in the department of which Professor McDougall was head; it was generally understood in those days that research in parapsychology was approved by the Department. J.G.P. was a graduate student in psychology, specially employed as research assistant to J.B.R. From the viewpoint of objectivity, it should be noted that J.G.P. had not at that time shown special interest in the problems of parapsychology, and in fact worked on other problems for his graduate researches. It was not until some years later that he decided to devote his energy to parapsychology.

The subject, Hubert E. Pearce, Jr., was at the time a student in the Divinity School at Duke. He had introduced himself to J.B.R. approximately eighteen months earlier and had stated that he believed he had inherited his mother's clairvoyant powers. In ESP card tests given by J.G.P. and J.B.R. during the intervening period he had exceeded the average score to be expected from chance in practically every experimental session under a wide variety of conditions. During that period he had participated in tests involving nearly 700 runs through the standard deck of ESP cards, averaging approximately 32% successes as compared with the mean chance expectation of 20%. Nothing like this prolonged series of tests had ever been made up to that time, and H.E.P.'s performance was recognized even then as highly exceptional.

The Distance Series was the first step involving different buildings in the separation of H.E.P. from the target card he was attempting to identify. The move was not so much a strictly necessary requirement for the exclusion of visual cues as it was a matter of providing a conspicuously wide margin of safety against the possibility of such cues. The use of different buildings, incidentally, was convenient for the independent recording of the subject's re-

sponses and the card sequences. It became easily possible at the same time to provide for duplicate recording and independent checking.

To those of us who had participated in the long series of earlier tests with H.E.P. under gradually improving conditions of test and observation, this further advance in experimental conditions was hardly required. The essential safeguards had already been approximated. There is, however, a tendency of the mind, when confronted with so incredible a hypothesis as that of ESP, to exaggerate the possibility of alternative factors such as visual cues, recording errors, the loss of records, and the like. The revolutionary character of the ESP hypothesis, then, made necessary a range of precautions that were not normally considered a part of the routine of experimental psychology. This atmosphere of critical apprehension concerning the adequacy of the design needs to be taken into account, for it was a part of the actua' situation in which the experiment was conducted.

Some idea of the state of mind prevailing at the time can be gained from the circumstances leading to the planning of Subseries D. Subseries A, B, and C had been designed on the assumption that no error was possible that could favor the ESP hypothesis—not unless the two men, J.G.P. and H.E.P., were deliberately to conspire to produce a fraudulent set of results. Wisely (and accurately) anticipating that there would be those who would find it easier to suspect collusion than to accept ESP as established, Professor McDougall recommended that J.B.R. identify himself with the actual performance of at least a short subseries of the distance tests in order that a theory of collusion would have to involve all three of the participants in the experiment. On the basis of this plan Subseries D was conducted with J.B.R. actively officiating with J.G.P.

Actually the primary research objective in the experiment was to compare the effect of short and long distance on the results. In the planning of the test series, this concern with the role of distance was the essentially novel feature of the experimental design. In most of the tests in which H.E.P. took part during the preceding period, the target cards had been within a yard of him. It was considered a sufficient first step to introduce a distance of at least a hundred times that unit as one that should reveal any effect of dis-

tance on any possible radiant energy that conceivably intermediated in the operation of ESP. Later in the series this distance was increased still farther. While, then, for the general public and the critic especially, the Pearce-Pratt Series came into focus as a conclusive demonstration of the *occurrence* of ESP, to the workers in the Parapsychology Laboratory it became the first definite step in the testing of the hypothesis of the *non-physical nature* of psi, the hypothesis suggested by earlier experimental work as well as by the study of spontaneous psi experiences.

Procedure

A single subject, H.E.P., was tested for his ability to identify ESP test cards manipulated by the experimental assistant, J.G.P., in another building, part of the time at a distance of 100 yards and part of the time at a distance of more than 250 yards from the location of the subject. The experiment was designed to test for the clairvoyant type of ESP; and J.G.P., accordingly, did not know the card order in the test.

Aside from planning the experiment, J.B.R. participated only in the independent checking of results, except for Series D in which he participated with J.G.P. as the witness to the operation of the test.

There were, in all, four subseries, A, B, C, and D, totaling 74 runs through the pack of 25 cards; and the series extended from August, 1933, into March, 1934. The testing days were not consecutive, though within a given subseries they were more or less so. They were selected, however, at the mutual convenience of H.E.P. and J.G.P. Subseries C was begun in October, 1933, and four runs were added to it in March, 1934, with Subseries D following thereafter. Specific dates may be found in Table 1. Subseries A was done with the 100 yards distance, Subseries B at 250 yards, and the other two subseries back at 100 yards. The 74 runs represent all the ESP tests made with H.E.P. during this experiment under the condition of working with the subject and target cards in different buildings. It was, in fact, the only distance test involving different buildings done at the Duke Laboratory at the time.

Series A was set up with an advance commitment on termination point. It was agreed that 300 trials were to be given H.E.P. The

following Subseries, B, was intended to be a duplication with only the additional distance involved, but the experimenters were interested in the big shift of scoring level from day to day which was shown at the longer distance. It was decided to allow H.E.P. to continue further so as to see what would happen. Subseries C was intended to be a repetition of Subseries A consisting of 300 trials designed to discover whether the lower scoring rate of Subseries B at the longer distance was a result of the altered situation or whether H.E.P. had declined in scoring ability. Subseries D, as has been stated, was intended as introducing a check on J.G.P., and its length was agreed upon in advance (150 trials, or six runs).

In actual operation the experiment proceeded as follows, regardless of which subseries was involved: At the time agreed upon, H.E.P. visited J.G.P. in his research room on the top floor of what is now the Social Science Building on the main Duke campus. The two men synchronized their watches and set an exact time for starting the test, allowing enough time for H.E.P. to cross the quadrangle to the Duke Library where he occupied a cubicle in the stacks at the back of the building. From his window J.G.P. could see H.E.P. enter the Library.

J.G.P. then selected a pack of ESP cards from several packs always available in the room. He gave this pack of cards a number of dovetail shuffles and a final cut, keeping them face-down throughout. He then placed the pack on the right-hand side of the table at which he was sitting. In the center of the table was a closed book on which it had been agreed with H.E.P. that the card for each trial would be placed. At the minute set for starting the test, J.G.P. lifted the top card from the inverted deck, placed it face-down on the book, and allowed it to remain there for approximately a full minute. At the beginning of the next minute this card was picked up with the left hand and laid, still face-down, on the left-hand side of the table, while with the right hand J.G.P. picked up the next card and put it on the book. At the end of the second minute, this card was placed on top of the one on the left and the next one was put on the book. In this way, at the rate of one card per minute, the entire pack of 25 cards went through the process of being isolated, one card at a time, on the book in the center of the table, where it was the target or stimulus object for that ESP trial.

In his cubicle in the Library, H.E.P. attempted to identify the target cards, minute by minute, and recorded his responses in pencil. At the end of the run, there was on most test days a rest period of five minutes before a second run followed in exactly the same way. H.E.P. made a duplicate of his call record, signed one copy, and sealed it in an envelope for J.B.R. Over in his room J.G.P. recorded the card order for the two decks used in the test as soon as the second run was finished. This record, too, was in duplicate, one copy of which was signed and sealed in an envelope for J.B.R. The two sealed records were delivered personally to J.B.R., most of the time before J.G.P. and H.E.P. compared their records and scored the number of successes. On the few occasions when J.G.P. and H.E.P. met and compared their unsealed duplicates before both of them had delivered their sealed records to J.B.R., the data could not have been changed without collusion, as J.G.P kept the results from the unsealed records and any discrepancy between them and J.B.R.'s results would have been noticed. In Subseries D, J.B.R. was on hand to receive the duplicates as the two other men met immediately after each session for the checkup.

Thus, from day to day as the experiment proceeded, H.E.P. was kept informed, as he had been in all his earlier experiments, as to the rate of success achieved. The practice of expressing enthusiastic congratulations should be mentioned as a part of the procedure. If, as rarely happened, the scoring rate was low, favorable emphasis was placed on the overall performance, the general average maintained, and the high standing of the subject in the comparative scale of ESP subjects. Throughout the series the paramount objective of high-order performance was held before the subject with all the vigor and expectation that could be communicated.

Results

General Evaluation

Since they were one series of tests carried out under essentially the same conditions, the four subseries (totaling 74 runs, or 1850 trials) may be pooled. Mean chance expectation is 20%, or 370 hits. The total number of successes actually scored for the series is 558, which is better than 30%. The theoretical standard deviation derived on a conservative basis is 17.57. This total of 558 hits

is 188 above the theoretical expectation and it gives a critical ratio of 10.70. The probability that a critical ratio so large as this would occur on the basis of random sampling is less than 10^{-22}. In the determination of the critical ratio given above, allowance is made for the slight correction applicable when, as in this experiment, the balanced ESP deck is used; that is, when there are five of each symbol in each pack. The variance of scores obtained with the 5 x 5 ESP deck depends upon the frequency with which the subject calls the different symbols. The largest variance results when the subject always calls exactly five of each symbol, and the SD of 17.57 was obtained on this assumption (**2**). However, the subject rarely called five of each symbol in a run, and the exact SD would therefore be smaller than the one used here, which makes the estimate of statistical significance a conservative one.

Table 1

PEARCE-PRATT DISTANCE SERIES: GENERAL RESULTS

Subseries	Dates		Runs	Dev.	SD	CR	P
	Start	End					
A.....	8/25/33	9/ 1/33	12	+59	7.07	8.35	$<10^{-14}$
B.....	9/ 2/33	9/30/33	44	+75	13.54	5.54	$<10^{-8}$
C.....	10/18/33	3/10/34	12	+28	7.07	3.96	.000075
D.....	3/12/34	3/13/34	6	+26	5.00	5.20	$<10^{-8}$
Total...	8/25/33	3/13/34	74	+188	17.57	10.70	$<10^{-22}$

Each of the four subseries is independently significant, as may be seen by reference to Table 1. The table shows for each subseries the date, number of runs, deviation, standard deviation, critical ratio, and the associated probability.[1] A complete record of the card order and calls for the series has been furnished from time to time to qualified workers who wish to make some special study of the material. This practice will continue.

[1] In the two reports, mentioned above, in which the run scores of the series were published, the scores of Subseries B and C were not given consecutively, and there were two other minor errors. It seems worth while, therefore, to list the complete run scores in chronological order here. The division between days or sessions is indicated by the use of semicolons.
SubseriesA: 3; 8, 5; 9, 10; 12, 11; 11, 12; 13, 13, 12.
Subseries B: 1, 4; 4, 4; 7, 6; 5, 0; 6, 3; 11, 9; 0, 6; 8, 6; 9, 4; 10, 6; 11, 9; 5, 12; 7, 7; 12, 10; 6, 3; 10, 10; 6, 12; 2, 6; 12, 12; 4, 4; 3, 0; 13, 10.
Subseries C: 9, 8; 4, 9; 11, 9; 5, 4; 9, 11; 2, 7.
Subseries D: 12, 3; 10, 11; 10, 10.

Results of Further Studies and Analyses

Since this series was first reported, the data have been used by a number of research workers at the Parapsychology Laboratory for additional analyses. Some of the analyses bear upon the general question of whether the target order was sufficiently random to justify the assumptions underlying the statistical methods used in the evaluation of the results. Other analyses were aimed at trying to discover further psychological information relevant to questions of how ESP operates. The following review includes most of these analyses, though it does not cover all of the critical reviews and discussions.

Tests of Assumptions Underlying Statistical Methods. Greenwood and Stuart (3) did a cross-check in which the subject's calls were arbitrarily matched against the cards of the third run following, the calls for the first run being checked against the card sequence of the third run, the calls of the second against the cards of the fourth, etc. To make the cross-check series the same length as the experimental series, the calls of the last two runs were checked against the cards of the first and second runs respectively. The 74 cross-check runs give a total of 387 hits, a deviation of 17 above mean chance expectation, which is less than one standard deviation.[2] The empirical variance of the cross-check run score distribution is 4.772, which is a close approximation to the expected value when the size of the sample is taken into account. Greenwood (2) also obtained the variance of the series by an exact method which takes into account the actual frequency with which the subject called the different symbols in each run. The average variance per run by this method was found to be 4.116, which is slightly greater than the theoretical variance of 4 for the binomial hypothesis which applies to the open deck or random order of cards and the value of 4.167 for the matching case based upon comparing two adequately shuffled closed decks. All four variances agree so closely in value that it makes no difference in the conclusions drawn from the data which of the four is used to compute the critical ratio. Stuart (10) calcu-

[2] The cross-check score as originally reported was 385 hits. In the present paper the practice followed has been to report the corrected figure when any analysis previously published has been found to have an error. All of these corrections are trivial and none affects any interpretation of the findings. The student who notices any such discrepancies should give this review precedence over the earlier publications.

lated the empirical variance from the run scores and found a value of 12.83. The extreme range of the scores, from 0 to 13, with the tendency for scoring to be below chance on many runs in Subseries B, contributed to the large empirical variance; but even so, the Pearce-Pratt Series is highly significant when evaluated by the empirical SD (CR=6.10, $P < 10^{-8}$).

Analyses for Secondary ESP Effects. The data of this experiment were checked for displacement by Russell (**9**), who compared each call with the targets in the run for as many places away as the position of the call permitted. In the usual terminology this means that for backward displacement the data were checked for -1 through -24 displacement; and for forward displacement, for $+1$ through $+24$. No evidence of displacement was found.

The Pearce-Pratt Series was included by Pratt (**5**) among a number of high-scoring experiments, the results of which were studied to see whether the hits were clustered or whether, conversely, they were distributed as if they were in a random series. There was no evidence of grouping of hits in the Pearce-Pratt Series nor in any of the other ESP and PK data analyzed.

This experiment was also included among those surveyed by Pratt (**4**) in his analysis of ESP data to determine if there was any evidence that the subject interrupted or changed his habitual sequence of calling after making a hit. The Pearce-Pratt Series did not yield any evidence of change-of-call; that is, there was no difference between the frequency with which the subject followed a response that made a hit by the different ones of the five symbols in his next call and the frequency with which one of his calls that made a miss was followed by the various symbols.

Also, without giving the detailed figures, we can report here that there was no evidence that could be detected by a chi-square analysis to indicate that the subject interrupted his habitual sequences of symbol association at the point of making a hit. This is a question that needs to be examined in longer series of high-scoring subjects where a weak effect would more likely be revealed by the statistical measures applied. If it is true that a subject scores well above the level attributable to chance without deviating from whatever habitual sequence preferences he may have, this fact might pro-

vide an important clue regarding the manner in which ESP impressions are brought to conscious expression.

In still another analysis of this series, Cadoret and Pratt (1) examined the misses in the subject's trials to see if there was any evidence of consistent wrong associations between responses and target symbols. No evidence of consistent missing in the Pearce-Pratt Series was found, though evidence in the results of other experiments led to the tentative conclusion that consistent missing was a genuine secondary effect.

All in all, the results of the analyses for secondary effects that have been made with the Pearce-Pratt data add up to a strong indication that H.E.P. was successful in achieving what he was attempting to do; namely, to direct his ESP calling upon the target for that trial, the card that was on the book at the moment the call was made. The single exception to this rule is one that was apparent while the series was still being done. That was the tendency in Subseries B for the subject to score below chance in some runs. That subseries produced a remarkable number of low scores, too many to be attributed to random fluctuation, though the above-chance scores still predominated to such an extent that the total score of the 44 runs of Subseries B is significantly above chance expectation.

Discussion

Viewed twenty years later, the results of the Pearce-Pratt Series still appear to allow no interpretation except that they were due to extrasensory perception. So far as the extrachance character of the series has been re-examined, it has led only to confirmation of the fact that the statistical significance of the results cannot be challenged. Moreover, no tenable sensory interpretation has even been proposed to explain the data. The recording-error hypothesis has presented no reasonable claim for support, and the additional studies of the card distribution in the series have shown no peculiarities of patterning that could in any way alter the conclusions. Even the collusion hypothesis would have to involve all three of the participants in a deliberate conspiracy. And finally, the scores were high enough and consistent enough over the series that common sense without benefit of involved mathematics would assure us that the series was not mere chance.

If, then, as the Warner survey revealed (**11**), these results, along with all the other researches on extrasensory perception, failed to establish a convincing case for ESP for the majority of the members of the American Psychological Association responding to Warner's questionnaire, it can safely be said that the issue is not a matter of scientific evidence. The series contributed all that an experiment can do toward establishing the ESP hypothesis. The rest is a question of receptivity on the part of the professional group.

It is true, the conclusions are circumscribed. Only one subject was involved and from this no generalization can be made regarding other subjects. Only a short span of time in a lifetime of the individual subject, H.E.P., was involved, and from this one study it would not be possible to make a general statement beyond those limits. This particular study could not claim to have brought out what enabled Pearce at this time to score at this percentage rate or to enable a prediction that he could, under definable circumstances, repeat this performance. None of these things, however, has been claimed for the Pearce-Pratt Series. It was enough to say that under the circumstances, within the limits, and at the time, an extrachance performance did take place that could not be ascribed to anything but the operation of extrasensory perception, whatever that is and however it is to be explained. But to have reached this point with the degree of reliability and unequivocality represented was sufficient to constitute a turning-point in the thought and experimental plans of the laboratory concerned. From that point on, another problem, another stage in the investigation of ESP was in order.

While incidentally intended to provide improved test conditions, the Pearce-Pratt Series introduced a comparison of distances. In the 30 runs of tests made at the distance of 100 yards in Subseries A, C, and D, the average per run was 8.8 hits. The average for all of H.E.P.'s 690 runs made up to the time the series began was 7.9 per 25. The comparison of the results indicated, then, that this order of distance could not be considered a limiting condition. In this experiment, at least, the subject's scoring could not be said to have been lowered by the introduction of distance at the 100-yard stage.

With the longer distance, 250 yards, introduced in Subseries B, the 44 runs averaged 6.7 hits per 25. This might, on the basis of averages, raise the question of a possible effect of the longer distance.

A closer look, however, at the chronological score distribution on page 171 discourages that type of interpretation. The score distribution suggests, rather, some other factor, something that apparently raised or lowered the scoring for a given day's work. Whatever the factor was, it was not one of mere distance between the subject and the pack of cards. At least, so the reasoning went at the time, and the experimental results were taken to suggest, though not conclusively to prove, that distance of this order was not a factor related in any essential way to the operation of ESP. Since the distances were comparatively short, the results suggested the importance of tests with longer distances. And the suggested absence of a relationship between ESP and distance logically raised the question of the relationship of ESP to time. This experiment thus had much to do with precipitating the experimental investigation of precognition, which was begun with the same subject, H.E.P.

Nothing stands out now in retrospect more strongly than the shift of importance of different features of such an experimental research as this one. During the intervening decades the Pearce-Pratt Series came to have its greatest value because of the character of the experimental conditions under which it was done. Its principal contribution to the understanding of the nature of ESP—its bearing on the question of distance—was hardly noticed, and it may still be some time before it is. However, it did accomplish its purpose in turning the attention of the workers immediately concerned to problems that logically followed acceptance of the results as showing no effect of the distance involved.

Today, however, for the experimenters who are confronted with the more urgent problems of the field and who look back at this series, the paramount feature is the exceptionally high scoring produced and held for so long a time by this individual subject, H.E.P. There had been in his earlier work ample evidence that H.E.P. worked up to a crescendo of enthusiasm and ambition in his tests and that in general his scoring rate followed the rise and fall of his motivation.

As already indicated, he actually rose above his earlier scoring when introduced, in the first subseries, to the distance test conditions. It was as if this was a special challenge to him, something of a climax, as indeed it was. In the general intellectual atmosphere

which he breathed, the conquest of distance was a climactic undertaking. The eye falls, too, on the striking shiftiness of scoring introduced when the distance was increased to 250 yards. It was shortly after this series was over, a matter of months, that H.E.P. suddenly lost his capacity to score significantly in the card tests under any of the conditions or in any of the types of tests in which he had performed so steadily and brilliantly for about two years. This special series may have a point of relevance to the primary problem of parapsychology today, the problem of acquiring control over psi performance.

References

1. Cadoret, R., and Pratt, J. G. The consistent missing effect in ESP. *J. Parapsychol.*, 1950, **14**, 244-56.
2. Greenwood, J. A. Variance of the ESP call series. *J. Parapsychol.*, 1938, **2**, 60-65.
3. Greenwood, J. A., and Stuart, C. E. Mathematical techniques used in ESP research. *J. Parapsychol.*, 1937, **1**, 206-26.
4. Pratt, J. G. Change of call in ESP tests. *J. Parapsychol.*, 1949, **13**, 225-46.
5. ———. Trial-by-trial grouping of success and failure in psi tests. *J. Parapsychol.*, 1947, **11**, 254-68.
6. Rhine, J. B. *Extrasensory Perception*. Boston: Boston Society for Psychic Research, 1934.
7. ———. Some basic experiments in extrasensory perception—a background. *J. Parapsychol.*, 1937, **1**, 70-80.
8. ———. Some selected experiments in extrasensory perception. *J. abnorm. soc. Psychol.*, 1936, **31**, 216-28.
9. Russell, W. Examination of ESP records for displacement effects. *J. Parapsychol.*, 1943, **7**, 104-17.
10. Stuart, C. E. In reply to the Willoughby "critique." *J. abnorm. soc. Psychol.*, 1935, **30**, 384-88.
11. Warner, L. A second survey of psychological opinion on ESP. *J. Parapsychol.*, 1952, **16**, 284-95.

Parapsychology Laboratory
Duke University
Durham, North Carolina

A REAPPRAISAL OF THE MATHEMATICAL EVALUATION OF THE REINFORCEMENT EFFECT

By T. N. E. Greville

History of the Problem

In 1936 Soal worked with two successful subjects whose extra-chance scoring came on the cards immediately preceding and immediately following the one at which a sender was looking when the subject made his call. He found suggestive evidence that this displacement scoring was higher when the preceding and following cards were of the same symbol than when they were different, as though the ESP impression from one displacement reinforced that from the other. This impression was confirmed in 1942-43 in a further series of tests with one of the same two subjects. In 1951 Pratt published an analysis of the evidence for the reinforcement effect provided by the work of Soal and Goldney, using two different evaluation procedures (2). One of these was suggested by A. M. Walker and has been described elsewhere by the present writer (1). In essence, it was derived on the assumptions that (1) ESP operated only on certain trials, and on these trials invariably produced a displacement hit (forward or backward or both); and (2) the results of all trials on which ESP did not operate were governed entirely by chance. This procedure has been criticized by Robbins, who proposed an alternative method of evaluation described by Pratt in the article already cited. Robbins' comments on the Walker-Greville method are reported by Pratt as follows:

> Dr. Robbins pointed out that the method of evaluation proposed by Mr. Walker was based upon the assumption that ESP operates to give true cognitions on some trials and that chance is effective on the remainder of the trials. While this assumption may seem to some people to be the most reasonable one to make, Dr. Robbins took the view that it is better not to introduce any unnecessary assumptions. He argued that there are a large number of different assumptions which might be made on *a priori* grounds. For example, it is possible to assume that on each trial the subject knows by ESP one symbol which is not the target and he calls one of the other four symbols with a ¼ probability of success. The use of any of the possible

assumptions would lead to a different formulation of the statistical problem and therefore to a different result. Such an approach is therefore a test of the underlying assumptions, as well as a test of the reinforcement effect. If a significant result is obtained under these circumstances, the reason may be that there is a reinforcement effect in the data, but it could just as well be that the underlying assumptions were incorrect.

Robbins therefore suggested a more direct approach to the evaluation of the reinforcement effect, simplifying the problem by considering only one direction of displacement. If it is decided, for example, to study forward displacement, we construct a fourfold table in which the non-overlapping target pairs designated for evaluation are classified (1) as forward displacement hits and misses and (2) according to whether the target cards preceding and following the selected call were of the same symbol (multiply-determined pattern) or different symbols (non-multiply-determined pattern). A chi-square test is then applied to the resulting table to test the hypothesis that the designated patterns came from a population in which the proportion of hits is the same for multiply-determined and non-multiply-determined patterns. The designation of the non-overlapping target pairs to be investigated and of the direction of displacement to be considered is made, of course, before tallying is begun.

Critique of Robbins' Proposal

It cannot be denied that Robbins' criticisms of the Walker-Greville method are well taken. Recent studies of "psi-missing" (3) make it abundantly clear that the operation of ESP is not by any means the "all-or-none" proposition envisaged in the assumptions underlying that method. In the present state of ignorance as to the precise mode of operation of psi processes, any method of evaluation which makes specific assumptions of this nature is less than satisfactory.

However, when Robbins' alternative proposal is carefully examined, it also is found to be open to serious objection as a procedure for evaluating a reinforcement effect. Even though we decide to confine our attention, for evaluation purposes, to forward displacement, the fact remains that the reinforcement effect is most likely to be observed when there has been positive scoring *on both forward and backward displacement*. Therefore, in comparing the

relative frequency of forward displacement hits in multiply-determined and non-multiply-determined patterns, we must reckon with the fact that in a multiply-determined pattern every forward displacement hit is necessarily also a backward displacement hit, while in a non-multiply-determined pattern every forward displacement hit is necessarily a backward displacement miss. Thus, a general tendency toward positive scoring on backward displacement, unaccompanied by any particular reinforcement effect, would automatically tend to produce a somewhat higher proportion of forward displacement hits in multiply-determined than in non-multiply-determined patterns.[1] Even by arbitrarily limiting the evaluation procedure to one direction of displacement, Robbins' proposal does not therefore succeed in avoiding entirely the difficulties inherent in simultaneous consideration of both directions of displacement.

Difficulties in Finding a Satisfactory Method of Evaluation

While the spectacular figures emerging from the work of Soal and Goldney strongly suggest, even on casual inspection, that something like "reinforcement" has occurred, a precise evaluation of the effect presents peculiar difficulties—greater, I think, tnan any of us realized when our examination of the problem was begun. This is because, on the one hand, it is desirable, as Robbins has so forcefully pointed out, to avoid making specific assumptions as to exactly how ESP operates in this situation; while, on the other hand, without making such assumptions it becomes most difficult to formulate a null hypothesis to the reinforcement effect: that is, to decide just what proportions of hits and misses are to be expected for multiply-determined and non-multiply-determined patterns in the presence of positive scoring on both forward and backward displacement but in the absence of reinforcement. This difficulty remains (for the reason previously stated in commenting on Robbins' alternative proposal) even if we consider hits and misses in relation to only

[1] This may be illustrated by referring to the probability model assumed in the writer's previous note. If p_1 and p_{-1} are the respective probabilities, on any given trial, of a precognition and a postcognition of the target, and the two are considered to be mutually exclusive, the probability of a forward displacement hit is $p_1 + \frac{1}{5}(1 - p_1 - p_{-1}) = \frac{1}{5}(1 + 4p_1 - p_{-1})$ in a non-multiply-determined pattern and $p_1 + p_{-1} + \frac{1}{5}(1 - p_1 - p_{-1}) = \frac{1}{5}(1 + 4p_1 + 4p_{-1})$ in a multiply-determined pattern. The latter probability exceeds the former by p_{-1}. The writer is indebted to A. M. Walker for this observation.

one direction of displacement. Multiply-determined and non-multiply-determined patterns involve different target structures, and the manner in which ESP will be expected to alter the chance proportions of hits in the two situations depends on the specific way in which ESP is assumed to operate. It is conceivable that psi-missing may operate in a large proportion of cases, and that the scope of operation of pure chance may be quite limited indeed.

An Empirical Approach to the Evaluation Problem

It seems to the writer that the best course to follow in this situation is to formulate a purely empirical definition of what constitutes "reinforcement," and then, if such "reinforcement" is demonstrated to have occurred, to try to ascertain, with the help of other types of analyses, what it can teach us about the mode of operation of ESP. We might, for example, say that, in the absence of reinforcement, if the proportion of (double) hits in multiply-determined patterns is $\frac{1}{5} + k$, the proportion of single hits (of both types combined) in non-multiply-determined patterns should be $\frac{2}{5} + k$. In other words, we say that the observed proportion of hits should deviate from chance expectation by the same amount in both cases. Any significant departure from this situation, in the direction of a greater proportion of hits for the multiply-determined patterns, will, by definition, constitute a reinforcement effect. As was done in the case of the Walker-Greville method, we could determine the value of k for which the likelihood of the observed numbers of hits and misses is greatest, and compute a chi-square to test the hypothesis that the observed data were taken at random from a population characterized by this value of k. Somewhat unhappily, it turns out that it is necessary to solve a cubic equation (rather than a quadratic, as in the case of the Walker-Greville method) to obtain the value of k, and it also appears that this equation is likely to have more than one root within the range of possible values of k, so that special care is required to select the root which corresponds to the maximum value of the likelihood function. Thus, the routine use of this method would involve serious computational problems for the ESP investigator.

Let us, therefore, take a second look at the Walker-Greville method, examining it from a purely empirical point of view, quite

apart from the questionable assumptions originally made in deriving it. The method is characterized by the fact that the expected proportions of hits are taken to be $\frac{1}{5} + (\frac{4}{5})p$ for the multiply-determined patterns and $\frac{2}{5} + (\frac{3}{5})p$ for the non-multiply-determined patterns, p being, of course, the same fraction in both cases. It will be noted that this would provide a more conservative definition of reinforcement than that suggested in the preceding paragraph, since the deviation from chance expectation of the proportion of hits among the multiply-determined patterns is permitted to exceed that in the non-multiply-determined patterns in the ratio of 4 to 3, before any tendency toward "reinforcement" is considered to be present. There is a certain inherent logic in this convention, since, in both types of patterns, the deviation above chance expectation is expected to be the same fraction of the total range between chance expectation and complete success. In other words, $p = 0$ corresponds to chance expectation, while $p = 1$ corresponds to the situation in which every trial produces a hit on forward or backward displacement or both. It is possible, therefore, to forget the reasoning by which the method was originally derived, and to regard it as implementing a purely empirical criterion for the occurrence of reinforcement.

Conclusions

While fully supporting Robbins' objections to the assumptions on which the original derivation of the Walker-Greville method was based, the writer finds a fatal objection to Robbins' alternative proposal in that it fails to distinguish between a genuine reinforcement effect and differences in the observed proportions of hits which would necessarily tend to result from above-chance scoring on both forward and backward displacement—the principal situation in which a reinforcement effect has been thought to occur. A re-examination of the Walker-Greville method suggests that it can be derived, without reference to the original assumptions very properly rejected by Robbins, from a purely empirical definition of "reinforcement," which is a conservative definition in the sense that it characterizes as "reinforcement" only a deviation from chance expectation of the observed proportion of hits which is substantially greater for multiply-determined than for non-multiply-determined patterns. The writer concludes, therefore, that the Robbins method is not a satisfactory

procedure for evaluating a reinforcement effect, but that the Walker-Greville method may be considered reasonably satisfactory until something better is proposed, or until the nature of the observed effect is clarified in such a way as to make some other type of analysis more appropriate.

References

1. Greville, T. N. E. A method of evaluating the reinforcement effect. *J. Parapsychol.*, 1951, **15**, 118-21.
2. Pratt, J. G. The reinforcement effect in ESP displacement. *J. Parapsychol.*, 1951, **15**, 103-17.
3. Rhine, J. B. The problem of psi-missing. *J. Parapsychol.*, 1952, **16**, 90-129.

EDITORIAL

RATIONAL ACCEPTABILITY OF THE CASE FOR PSI

In the establishment of scientific discoveries in natural science there are certain discernible stages. Even though the process as a whole may well be one continuous development from original question to accepted conclusion, the character of the procedure changes sufficiently in its course that it can profitably be subdivided for discussion and better understanding. In a preceding editorial we have already discussed two broad divisions of this general procedure of scientific investigation. The first of these was the exploratory stage characterized by the search for trial answers or working hypotheses to solve the problem in hand. The second stage was identified with the devices of testing the reliability of the working hypotheses developed in the earlier probings. It was shown that this verification division of the larger program of research is featured by the effort to design crucial experiments, to establish and confirm factual relations against the possibility of all alternatives relative to the problem.

At the end of the discussion referred to, a third stage in the scientific acquisition of knowledge was introduced. This stage involves the task of clothing the new discovery with a proper and presentable rationale adequate to making it acceptable to scientists in general. To leave a newly found fact on the doorstep of science without any evidence of the relationship it may have to already familiar knowledge of Nature is not enough at this stage of confidence in scientific method to win consideration for it by the average scientist. He is likely to regard the stranger as illegitimate and to use one of the numerous devices of the mind for avoiding recognition of it.

The preceding stage of verification should, it is true, develop an adequately controlled and really crucial experiment; and when, in addition to this, there is also independent confirmation of such an experiment, the confirmed finding should ideally be regarded by all as an established fact. But when a basically new advance is made that comes into conflict with prevailing theory, a further step of

rational confirmation is always needed for eventual acceptance of the conclusions.

The reason for this third stage in the establishment of a fact is that the whole procedure of science is made up of a series of observations, judgments, and interpretations that are human and potentially subject to error. It is very reassuring, therefore, if the new result is consistent with logical expectation. On the other hand, the more the new conclusion seems to contradict current thinking, the more important and prolonged becomes the third stage of making it rationally understandable. While experimental confirmation eliminates most of the uncertainty associated with the personal conduct of the research concerned, it still leaves to rational synthesis the need to remove whatever intellectual blocks confront the new conclusions.

It is a function of the problem-solving operation of science, then, not merely to introduce and verify a new fact but to install it, as it were, in the rational setting in which it must work harmoniously with the other facts of Nature to which it is related. This does not mean that the new fact must conform to the old beliefs, opinions, and untested doctrines of the professional group concerned. To require the squaring of empirical facts with existing beliefs is a rank violation of the scientific way of inquiry. It is, however, a very different matter to buttress the new discovery with a supporting rationale of internal self-consistency and to stabilize its unsteady novelty with interrelationships with the groundwork of surrounding knowledge.

* * * *

Any rational approach to the phenomenon of psi should first take account of the important historical fact that the psi researches arose out of no mere speculative claim promoted by a professional group or philosophical cult. The psi hypothesis grew out of reports of spontaneous human experiences coming from various cultures and from a wide range of types of individuals within a culture. These personal experiences of psi, when collected and classified, clearly justified serious consideration and study. In almost any other field of research such evidence would be given great weight and credibility. The present argument is that scientifically psi has a much stronger case because it rests on the experimental verification of a hypothesis derived from the study of widespread reports of human

behavior than it would have had if a purely speculative invention had been the subject of investigation.

Parapsychology is, in fact, unique in the dual character of its method of investigation; that is, in the degree to which it has developed from parallel and mutually supporting lines of evidence—the case study line on the one hand and experimental tests on the other. It has thus been possible to make comparisons back and forth from one method to the other and acquire an added measure of reassurance from this type of confirmation. The same types of psi phenomena, for example, that have been verified experimentally are precisely those which have been manifested in the collections of case reports. Thus telepathy, clairvoyance, precognition, and psychokinesis as types of psi occurrences have a double order of support.

The case studies and the experimental tests have also supported each other with regard to some of the general characteristics of the psi phenomena. Both, for example, show independence of psi from time and space, both indicate unconsciousness of the psi operation, and both suggest a similar relationship of psi to motivation. For the most part, it is true, the use of case studies has been limited to the exploratory stage of the investigation. It was mainly from them that the hypotheses originated which later received experimental confirmation, but this will not likely be as much so in the future as the experimental work advances.

The experimental evidence itself divides easily into a number of main types, and this further diversification of the supporting material introduces new angles of relationship and very considerably reinforces the total impressiveness of the case. (All this is, of course, in addition to the great accumulation of just plain evidence of high scoring in the tests.) For example, it turns out to be quite common in psi experiments that the scoring rate declines in the run or in some other unit of the subject's performance. Statistical examination of these declines shows extrachance effects produced in this declining tendency alone. In other words, here is a secondary effect supporting the primary one, a phenomenon within the principal evidence of the phenomenon. These declines, moreover, lend themselves to rational explanation quite as well as do those of general psychology (memory curves, learning gradients, and the like) though the explanation is not necessarily the same.

When by the evidence a secondary discovery concerning psi is forced upon the experimenter and when it comes as a complete surprise, it has added value in making the phenomenon seem more objective and convincing. We may take the psi-missing tendency as an example. This tendency of some subjects under certain conditions to miss the target to a significant degree came as a wholly unexpected experimental result. It was, however, not difficult to find plausible explanations for such a tendency once it had been recognized and especially once it was also realized that subjects are wholly unconscious of the functioning of the psi process. It was understandable that a subject could direct his effort toward the card sequence serving as the target order without always being able to exercise enough control to keep from consistently giving the wrong responses to the target cards. In such a case, the more effectively he used psi the more consistently he missed the target. This curious evidence of "psi perversity" peculiarly well meets the rational need felt by some skeptics for test results to counteract the possibility that motivated errors on the part of the experimenter could account for the results.

Other variations of this unconsciously misguided aim of the psi function could be listed. The most conspicuous evidence of this is to be noted in the work of Soal and Goldney on displacement; that is, on the tendency of the subject to give a response corresponding to the target preceding or following the one intended. The consistent way in which this displacement occurred—for example, the fact that it was limited to one direction (forward or backward) with a given agent or sender and the fact that with one and only one sender did it spread both to forward and backward directions—gives a distinctly lawful appearance to the aberration. A number of other features and variations extend this concept of the orderliness in the working of ESP in the tests. No better example of this could be given than the regular way in which displacement, in the Soal and Goldney experiment just mentioned, extended to the second card ahead and the second behind the intended target when the speed of testing was accelerated. Such variation of performance with test condition provides a high grade of rational confirmation of the experimental results. The displacement effect, too, was as much of a

surprise to the experimenter as it was unintentional on the part of the subject.

It adds a great deal to the reasonableness of the case for psi that there have been a number of persistent individual peculiarities of performance which were characteristic of special subjects. The peculiar limitation of the displacement effect to certain individuals constitutes a case in point. On this principle of individual differences it should be expected that a difference great enough to show significance should be found in scoring rates between the different personality classifications if the personality characteristics of the subject could be measured and correlated with performance in psi tests. Accordingly, it has now been well established that certain roughly measurable attitudes and personality traits and states have some connection with the subject's performance in the tests. It does appear that these personality correlates largely exert their effect, not in determining how long at a time or how effectively psi will function, but rather in determining whether or not a subject's response will be one of psi-hitting or psi-missing. At least there is evidence of a significant relation between personality measurements and success in psi tests. Moreover, this evidence indicates that psi interoperates in some way and degree with the more familiar elements of personality. This evidence adds a firmer rational quality to the case for psi as well as a better understanding of its role in the total personality.

Perhaps for many inquirers who look into parapsychology the effect of introducing interferences into the test situation will carry a supplementary measure of rational conviction. The drug experiments (for example, those with sodium amytal) will be especially to the point, but of closely similar bearing are those in which distracting psychological conditions were introduced, such as the effect of boredom or deliberate distraction of attention. It should be no less important, however, to consider along with these interfering states all that is known about the conditions especially favorable to high scoring, such as the use of rewards and playful competition.

Along with the evidence of individual peculiarities and the special conditions, there are certain more general features to be considered as a part of the rational concept of psi. Some of these, such as the unconsciousness of the function, we have already mentioned.

The very consistency of this characteristic is remarkable, particularly if one considers the frequency of exceptions to generalizations about nature, especially in psychology. Psi has remained thus far an irrecoverably unconscious process and the research has revealed no exceptions. There is the same extraordinary consistency, too, in the evidence of the nonphysical character of psi. There simply has not been found any limiting relationship between the physical conditions and the test performance. This nonphysical quality becomes, then, a defining feature and the unexceptional lawfulness indicates that the boundary is no mere arbitrary one, as so many academic divisions of nature are. There is, therefore, a comparatively deep and important division between the phenomena on the two sides of the boundary. This distinction must be made, however, without any implication that the division is an absolute one.

The spontaneous, elusive, and uncontrolled character of the psi function is another of its most consistent features. In fact, this is probably more invariable than has been realized generally among parapsychology students themselves; it has been a difficult fact to accept. It is doubtful, however, if within the period of recorded history of the human species psi has ever been anything but elusive and uncontrolled. If any individual ever lived who could control psi, he surely would have been so elevated above the level of his fellows that a record of his unusual life story would have become a familiar part of human history. Not even any of the great religious leaders lived lives that suggest anything approaching an unlimited control of psi capacity. As has often been pointed out, however, the suggestion is that they may either have had an unusual range of exercise of it, or a special ability to use it wisely and to advantage.

It adds greatly to one's appreciation of psi to know the weakness of its real antagonist, the philosophy of materialism. The situation is actually not a little preposterous since that philosophy is based on no factual evidence whatever. There has, of course, never been any crucial experimental test of the hypothesis of materialism other than the experiments of parapsychology and they have conclusively disproved it. Consequently there is no evidence for the other side of the issue which the psi hypothesis would have to refute. The case

for materialism simply collapses in the face of the evidence against it from parapsychology.

The opposition of common sense to the case for psi is on similarly shaky ground, even in confronting that most puzzling problem, precognition. The fact is, practically all of the main discoveries of science have been in violation of the common sense of their day. Common sense catches up only as new discoveries themselves become converted into general knowledge and opinion. It is true, if causation properly applies to psi at all, the researches have produced results that seem to call for causation of a wholly new kind. In precognition it would appear that the order of causation is reversed and the perception precedes the event. This is not the first instance, of course, in which science has produced conflicts that seemed at the time to be unresolvable. The proper procedure in such cases is to suspend judgment until further knowledge is acquired by which to see the problem in a setting that will permit understanding.

The amount of internal consistency within the evidence for psi is of great importance to any appraisal of its entire value or meaning. On this aspect the case is especially strong. There has been found, for instance, no peculiarity of ESP of the telepathic type (except individual subject preferences and differences) that does not hold for clairvoyant ESP as well—that is, beyond the mere difference of targets, which distinguishes the two phenomena. Nothing has been discovered about precognition except the aspect of time that is not correspondingly true of contemporaneous ESP. Even in PK, except for a lower level of scoring rate and the nature of the test itself, there is nothing essentially peculiar. So the parapsychologist is naturally led to the working hypothesis that there is but one basic psi function underlying all the effects. It is even difficult and almost impossible sometimes to distinguish sharply and finally between these aspects of the manifestation of psi, so well-integrated is all of the experimental evidence. Neither do the various kinds of secondary evidence show any distinction between types of psi. All stand equally clear of physical involvement, all are equally unconscious, and all require subject motivation in a similar degree. The psychological conditions affecting one kind of psi affect all others similarly. This internal harmony of relationship is shown already even at this early stage of development of the field.

* * * *

The supreme rational test for a new discovery consists of the success attainable in fitting it into the structure of related knowledge. Psi has passed this test as far as it has been possible to apply it. There is no real conflict between psi and the science of physics simply because physics has no franchise on the kind of phenomena with which parapsychology deals. Physicists, fortunately, have been the first to recognize this fact. That psi phenomena and physical phenomena are both a part of a larger universe it is hardly necessary even to state. If there were not some larger unity, one could not conceive how there could be interaction and the interconversion of causal influence that is involved in the subject-object reaction with which experiments in parapsychology deal.

It is entirely conceivable therefore that new and as yet unsuspected ways of looking at the relationship between physical and psi processes will have to be found. It is possible even that current ways of reasoning may have to be amended. It seems wise, in the meantime, to continue with the logical course that has been most useful in the study of nature at least until it is clearly necessary to leave it; and thus far in dealing with psi phenomena it has been rationally possible to use the concept of causality most of the way. Doing so requires the assumption for the time being of a cause, an active principle or energy, operating between subject and object, even in the extremely puzzling case of precognition. The supposition of a psychic energy, however, has often been made by general psychologists before, but until parapsychology demonstrated the presence of an extraphysical function in man, the assumption of a mental energy was never a *necessary* one. The materialist could always argue that perhaps some combination of the physical energies of the brain would suffice. Now it is known that the hypothesis of a mental energy is required.

The reasoning at this point must necessarily still be speculative, but it is pointing a way that empirical test can and will have to follow. Moreover, once a causal but nonphysical operation underlying psi effects is assumed and once a psychic energy is inferred, a theoretical basis for a tentative discussion of the place of psi phenomena in the natural order is available. Psi evidently is a mediating, energetic operation between the subject and his environment, one that supplements the sensorimotor system. Assuming a causal

relation then, it is logical to invoke the Law of Reaction and, given the fact of a subjective perception of an object by ESP, to "predict" the psychokinetic effect upon an object by a properly motivated subject. And that effect the PK research has demonstrated in the case of moving objects (dice, etc.). Thus, even while leaving some first-rate problems still to be solved, the integration of ESP and PK as phenomenal aspects of an underlying causal subject-object relation is made considerably more understandable by the psychic energy hypothesis.

For the rational critic of psi, the most difficult phenomenon to cope with is that of precognition. It will help a great deal to remember, however, that natural science grew up on the basis of sense data and their derivatives, and those data are sensorially mediated experiences of the physical world. In the very nature of the case the laws of physics limit the operation of sense perception; and in this circular nature of the whole development of science the scientist has fallen into the habit of thinking that the world derived from sense data represents the total universe. That world is defined in terms of time, space, and mass. Sense data would have to depend on energies thus conditioned; ESP data would not. Precognition is, therefore, a *reasonable* aspect of psi function, just as it seems quite unreasonable to the physicalistic habit of mind which all students of science have acquired.

When we attempt to integrate psi with general psychology there is only one major distinction to be made. Psi, in both its ESP and PK aspects, is fundamentally different from the sensorimotor system of the organism in the fact just mentioned: that the sensorimotor functions are physically limited and intermediated. The operations of ESP and PK are not. Psi phenomena are unconscious. Sensory functions are primarily conscious, at least in man, and though often only marginally conscious or even subconscious, they are apparently never unrecallably unconscious. ESP, on the other hand, seems to be unconscious beyond recovery. There are, of course, many other sensorially imperceptible and unconscious operations in nature, some of them that actually affect the organism. Thus far all that are known are recognizedly physical, but possibly no other kind would have been recognized.

If, now, we assume the existence of a psychic energy that is not

subject to the limitations of space, time, and mass, it is logical to assume further that unlike the known physical energies, it would not be sensorially perceptible or capable of producing directly conscious effects. It should be expected, rather, to have its own distinctly peculiar principles of operation and conditions of functioning. We may well inquire about the properties and functions of such an energy, how it interconverts, for example, to effects that are capable of crossing the threshold of the consciousness of the subject, and how it interoperates with an object that is the target of ESP or PK. But these are questions for the experimental researches of the future, for general psychology, for physics, for neurology, as well as for parapsychology.

The greater share of the mystery of the nature of psi is still left, to be sure. So far as showing consistency in what is already known is concerned, however, parapsychology does comparatively well—as well, it is safe to say, as any other frontier branch of the natural sciences. If psi is not inhibited by material barriers, if a thought as well as an object will serve as the target, if a present object as well as a future one can serve as the target to be perceived, and if an object can be perceived as well as influenced in its fall, then all of these operations at least take place within a consistent rational framework of nonphysical causation. Since they do not follow physical law they could not be expected to be sensorially perceived; if they do not follow spatial limitations they ought not to be expected to follow time; but since they do interoperate with the physical system they can be studied not only by objective methods but with quantitative measures as well.

More than half of the rational difficulty of the average student of science facing so new a phenomenon as psi is caused by sheer unfamiliarity with an unconventional concept. This effect of familiarity can best be appreciated if one starts inquiring into the explanation of his own simplest everyday experiences. Let him begin with the moment he awakes in the morning. Let him ask himself, what does it mean that he is awake? What really is sleep itself? Not physiologically but psychologically. If he answers that it means he is conscious, he is only rephrasing the question. He may go on describing his mental activity as that which constitutes being conscious. His words, however, are mere translation. He can never

get far beyond the initial unanalyzed fact that he is conscious and knows it. He will remain unable to put it in any more elemental or original explanatory terms. Such a rational inquirer, then, coming back to a psi phenomenon, can comfort himself that familiarity, which is at least partially attainable through study and practice, will dissipate a large share of what appears at first to be a more serious difficulty.

It prepares the anxious skeptic very considerably for a better appreciation of the evidence for psi to recognize that, after two thousand years of the philosophy of materialism and a good century and a half of the popular promotion of it, it has still not been deliberately and systematically applied to social institutions, except through the arbitrary force of a communist dictatorship. It is valid to argue that if there be an essential truth in a philosophy, then to allow it to flow into the educational currents of the time should permit it to find its way to justification through use. But in spite of the preponderance of materialistic thinking in connection with the science of the last two centuries, this philosophy has never taken hold in any culture except through enforcement by military power and political intrigue. Thus, in blocking the acceptance of the psi discoveries as it has, this philosophy of physical monism has not had the support either of experimental fact or of practical experience.

In brief, then, psi appears to have a strong, supporting rationale. Its only opposition comes from an untested assumption. It is based on a system of evidence of varied, complex, and consistent character. Its lawfulness has been revealed in many facets of its study, showing interrelationships that reinforce its reliability. It supplements, rather than conflicts with, the science of physics; and with an inferred psychic energy to account for its effects, it integrates with the nature of the universe and of man with no difficulties which would seem insurmountable to future research. The essential mystery of psi, which should not be minimized, is the great mystery of all mental life as such. What, indeed, is this conscious (and unconscious) willing, thinking, feeling system of operations, this mind that is pursuing the question of psi itself? Whatever that mind may, in all its rational ramifications, turn out to be, psi is an integral part of it firmly united now by the strong network of interrelationships that makes a science sound and reassuring.

J. B. R.

LETTERS AND NOTES

Dear Dr. Rhine:

It is agreed that analysis of the scoring decline rate for PK and ESP tests indicates a very close correlation between the two tests so far as mental conditions are concerned. Certain physical aspects such as variations of distance between subject and object, or of size of object are apparently not critical in either type of psi operation. There may, nevertheless, be physical forces which tend to level out the results of a PK test which are not present in ESP tests.

In the ESP card test, for example, the subject has the same chance to identify correctly a target card with a circle as he would have if the target happened to be a square. Consider, on the other hand, the use of dice in an effort to physically produce a given face. No accepted research to date has indicated that a subject can cause a die at rest to turn over and show the target face. Assume, then, that the die must be in motion. Obviously, if the motion should be fully constrained to a fixed pattern, the subject would have no more chance than before. It is quite possible that in a random throw the initial forces exerted through impact could be such that the outcome would be too strongly influenced (that is, the object would have too much inertia) to permit a PK force to affect it appreciably. On the other hand, it might be that at some point in the roll a point of unstable equilibrium is attained so that PK influence could cause the die to produce the target face. It appears only logical that the greater the number of such "crossroads," the higher the possible score of a good subject. Also, a question may exist as to whether there are not different types of crossroads, some more effective than others.

It would appear, then, that the number of hits produced in a dice test does not necessarily indicate the potential ability of the subject, since the number of hits above chance actually possible through the use of a psi capacity is unknown for any run.

The foregoing hypothesis could be tested, using a PK device in which the number of crossroads could be varied from one test to the next without seriously influencing the subject's interest or confi-

dence. Obviously, such a test will be difficult to formulate. An attempt at a suitable device has been made using glass balls, 5/8 of an inch in diameter, which roll down an incline over #4 grit sandpaper mounted on glass. The subject's task is to cause the balls to roll into a designated "target" side of the area. The use of different sizes of grits glued to interchangeable glass plates, together with the use of balls of different size or material, and a variable incline angle should provide the desired variation. Whether the type of crossroads thus provided will be effective is unknown. No tests have been made to date, but it is hoped that some exploratory results will be available soon.

Very truly yours,
Leslie C. Wilbur

(*Professor Wilbur is on the faculty of the College of Engineering of Duke University.*)

Dear Prof. Wilbur:

You have raised a very timely and pertinent question about PK testing with dice. Better still, you have proceeded to try to test your suggested hypothesis. If in a series of PK placement tests you can compare balls rolling over surfaces with an increasing order of roughness from one test condition to another, you should thus be able to evaluate your hypothesis. You are aware of the need to try to maintain a mental attitude free from bias toward the different conditions compared, though of course it is not easy to do.

You should be able to test your hypothesis also by comparing different shapes of dice. To do so you might begin with regular cubes for the first stage and shift in later comparisons to dice with progressively more rounded corners. The more rounded they become the more unchecked the roll down a smooth incline will be. Thus, the more nearly cubical the dice, according to your view (with which I agree), the more readily influenced they should be. It would be ideal to test the hypothesis by way of both approaches in the course of time. The result would be most revealing if it showed a limitation of PK effect tied in reliably with a physical variable.

Sincerely yours,
J. B. Rhine

* * * *

Dear Dr. Rhine:

Your editorial in the March *Journal* on "Perspective in Transverse" raises two points about which some disagreement—or perhaps difference in emphasis—may arise. These two points, as I gather them from this editorial and from other writings of yours, may be summarized as follows:

> 1. That schools of parapsychology might retard progress, and might mean a tragic waste of energy.
> 2. That survival is now an untestable, speculative hypothesis; that no one can propose a good research plan for a crucial test of it; that conclusive establishment of the hypothesis of survival by means of spontaneous cases is impossible; that the question of spirit survival has come to a "stalemate"; that the survival hypothesis may be "allowed to survive, but [that] it must be laid on the parapsychological shelf along with the telepathy hypothesis," there to "wait its turn" until far-sighted parapsychologists, by discovering whether any immaterial part exists in the human personality, provide leverage for competent treatment of the problem.

Before expressing divergent views on these two issues, it may be pertinent to state certain agreements which (the present writer believes) almost all thoughtful readers of the *Journal* support. First, that it *is* unscientific to endow "untestable hypotheses" with "more belief and significance than they deserve." Second, it *would* be unscientific for "easy speculative theorizing" to "gain ascendancy over experimentation." Third, parapsychology *does* have "to work out its own standards of evidence as well as its methods." Fourth, it *would* be a tragic waste of energy if schools of opinion were to develop among psychical researchers on the basis of such unscientific methods as those just cited.

Bearing in mind the above agreements, the undersigned would like to place on record the following propositions, not in a spirit of antagonism, but rather with a view to carrying out the ideals which you have so often advocated—namely, the well-rounded presentation of points of view held sincerely by thinkers who have given intensive attention and work to these problems:

> 1. Intellectual competition—and even conflict—can be highly stimulating when intellectual cooperation is unattainable. To suppress and resist fresh and vigorous points of view would be dogmatism.

2. Conclusive proof is always impossible—outside of pure mathematics. Parapsychological experiments create only probabilities.
3. Systematic and critical analysis of evidential spontaneous cases can develop fairly high probabilities for or against survival. As Professor Ducasse and others have pointed out, such rigorous analysis of spontaneous phenomena is consistent with the highest traditions of scientific research.
4. From such systematic analysis of spontaneous cases, promising hypotheses relative to survival—and sound experimental designs for testing them rigorously—may be developed.
5. While "easy speculative theorizing" may become a vice, certain basic philosophical problems must be confronted if the scientific study of spontaneous cases and the conduct of laboratory experimentation are to be fruitful in solving the problems of parapsychology. Among such problems are the nature of personality, and its relations to time, space, matter, and causation. To take for granted superficial and naïve assumptions on these basic questions would be just as grave a scientific error as to devote one's entire time and energy to untested philosophizing about them.

Sincerely and cordially,
Hornell Hart

(*Dr. Hart is Professor of Sociology at Duke University and is a member of the Advisory Editorial Board of this* JOURNAL.)

Dear Professor Hart:

I welcome your discussion of the editorial "Perspective in Transverse," and with your propositions for general agreement I am in full accord. The degree of my concurrence in your set of special proposals is, I feel, an unimportant individual matter. In any case my reservations are minor.

The survival hypothesis was, of course, mentioned only incidentally in the editorial. Your summary of my attitude, drawn from other sources, does not, I think, fairly represent the constructive effort I have spent on that problem over a period of more than thirty years; but since I have stated my present position at least briefly in my book *New World of the Mind,* I do not think I should repeat it here. Those of us who have a strong common interest in the problem, however, ought not to be expected to agree very closely on methods and appraisals of results at this exploratory stage.

Sincerely yours,
J. B. Rhine

BOOK REVIEWS

APPARITIONS. By G. N. M. Tyrrell. New York: Pantheon Books, 1953. 172 pp.

G. N. M. Tyrrell's theory of apparitions is already familiar to students of parapsychology. It was the subject of his Myers Memorial Lecture to the Society for Psychical Research which was published by the Society in 1942 under the title "Apparitions." This earlier edition has recently been edited by Mr. Edward Osborn and re-issued for a less restricted public. Accordingly, the original introductory chapter, largely directed to that earlier audience, has been omitted and a preface by Prof. H. H. Price, the Oxford philosopher, has been added. Except for minor editorial changes, however, the text devoted to the theory of apparitions is unaltered.

The preface increases the popular appeal of the book, not only on account of the prestige of the writer, Prof. Price, but also because his brief outline of Tyrrell's main theory will be a help, especially to general readers. It is needed, not because the text lacked clarity, but on account of the unfamiliarity of many such readers with the issues and concepts involved and even, in some instances, with the theme of the study.

Tyrrell's aim, according to Prof. Price, was "to reformulate the telepathic theory of apparitions . . ."; his interest was "nothing more nor less than the exploration of human personality and particularly the exploration of its subliminal or unconscious strata." For Tyrrell thought that "the chief importance of apparitions, as of other supernormal occurrences, was the light they might be expected to throw on the structure of human personality."

With this objective, then, Tyrrell examined cases of reported apparitional experience, using as illustrative examples sixty-one instances drawn from the older literature (*Proc. of the S. P. R.;* Gurney, Myers, and Podmore's *Phantasms of the Living;* and Myers' *Human Personality and Its Survival of Bodily Death*). The theories he attempted to reformulate were mainly those of Myers and of Gurney who, in an earlier day, based theirs on many of the same cases now restudied by Tyrrell.

It will be recalled that Myers and Gurney, being contemporaries, had argued out their respective concepts of the nature of apparitions, agreeing on some points and continuing to differ on others. Both had rejected the naïve idea that an apparition indicates an objective reality "out there." They agreed, too, that apparitions of the living are essentially telepathic experiences. They also agreed in their assumption as to the nature of the telepathic process. The assumption was that in a telepathic exchange the minds of two separated individuals were somehow in effective contact. One of the two (and this is a point to watch) was designated as the agent. He originated a thought or idea and managed somehow to "send" it to the other, the receiver or percipient, whose role was a relatively passive one.

The differences between the ideas of Myers and Gurney arose when they tried to account for instances of collectively seen apparitions. Myers found it necessary to postulate some sort of "psychical" (metetherial) but not physical invasion of the space sensorially present to the percipient to account for the fact that other observers besides the percipient sometimes share the apparition. Gurney preferred to suppose instead a type of telepathic contagion by which the experience of the percipient could be shared by others near him.

Tyrrell's theory advanced beyond these of his predecessors mainly because, by introducing a new concept of the unconscious, he overcame some of their difficulties. Both Myers and Gurney appreciated the fact that the mental action involved in psychic experiences is not conscious, and Myers' study of the subliminal is a classic. But Tyrrell, having the advantage of the thinking of the intervening years, was able to be more definite than the earlier psychic researchers could have been in formulating a concept of personality structure involving *levels* or layers of the unconscious. Certain of the layers "mid-level" between the conscious self and the physical body, then, become for Tyrrell areas critically involved in the production of apparitions.

At the same time, although less definitely, Tyrrell postulates, in Price's words, a "non-separate" character to the mid-levels of different personalities. "According to Tyrrell," Price says, "the word 'telepathy' is just a way of referring to the *non-separate* character which the mid-levels of different personalities have." According to

this concept, the physical distance between agent and percipient ceases to function as such.

But still Tyrrell, like Myers and Gurney before him, holds that the agent is the individual who must activate the process, even though he does not consider the agent responsible for getting a message or information "sent" to the percipient. Thus the agent's role is decreased with Tyrrell's non-separate mid-level concept, and at the same time the role of the percipient has been increased, for the percipient is no longer merely passively receptive. He now becomes the dramatist with the responsibility of working up a convincing effect for his own consciousness, either doing it by himself or, if necessary, with the cooperation of other individuals who may be present.

Thus, by eliminating "distance," Tyrrell gets around the old difficulty of explaining how in a telepathic process an agent could "send" a message across intervening space to the percipient. By the same means he resolves the troublesome difficulty of explaining collectively seen apparitions, since according to the hypothesis, the mid-level unconscious constituents of the bystanders are non-separate from those of the agent just as much as are those of the percipient himself.

It is of course impossible in a review to do full justice to theories as complex as these of Myers, Gurney, and Tyrrell, but the point must be stressed that all three hypotheses are based on the same theory of telepathy. Although Tyrrell modifies it considerably by giving much more of the action to the percipient than did the earlier writers, it is still in essence the old concept of mind-to-mind transfer, with the agent still credited with the important function of initiation.

How adequate, then, is Tyrrell's theory? A first requirement of any theory is that it be able to cover all the pertinent facts. There are likely to be few who will question that Tyrrell's theory logically can cover the portion of all apparitional data from which he derived it. But apparitions form only one segment (and a small and relatively obscure one at that) of all telepathic experience. However, without attempting to evaluate its applicability to forms of telepathy that are not apparitional (or to psi apparitions that are not telepathic), it will be sufficient to raise the question asked by Prof. Price when he says, after discussing various types of apparitions: "The

telepathic explanation does seem to fit some of the facts very well. Can it be extended to cover all of them? Most students of Psychical Research have assumed as a working hypothesis at least that it can. Tyrrell shared that belief."

But this telepathic explanation, historically an initial one, was based only on the superficial aspects of "thought-transference" situations. Even as Tyrrell wrote, it was already open to serious question. Even then, experimental evidence was suggesting that the basic ESP process may be the same whether the "target" be the thought of a human being, as in telepathy, or an inanimate thing, as in clairvoyance. Neither then, nor since, have any data been discovered to justify the idea that telepathy is a distinctive extrasensory process. It must not be forgotten that because of the still obscure status of the thought-brain relationship, there is no evidence that a direct mind-to-mind transfer in the original sense does occur. But if no basic distinction can be drawn, except speculatively, between two-person (telepathic) and person-thing (clairvoyant) situations, then an adequate hypothesis must explain both types of ESP. It will have to be different from the old telepathy hypothesis since in clairvoyance there is no *agent*. Therefore a theory to cover both can have only one necessary and active participant instead of two. For even though in some instances the agent appears to be active, the same process can go on when he is absent. The difference between a process involving an agent as Tyrrell conceived it, and one with no necessary agent at all may seem to some a slight one, but it is in reality fundamental. Radically different lines of thought arise from the two concepts.

Apparently, without realizing it, Tyrrell seems to have approached the no-agent idea closely when he found it necessary to assign the dramatization involved in the production of apparitions to the percipients instead of the agents, as the earlier writers had done. He approaches it also when drawing an analogy between sense perception and the seeing of apparitions. Certainly sense perception as illustrated by the seeing of an inanimate object (a brick) is more closely analogous to clairvoyance than to telepathy as he conceived it. Nevertheless, he skirted recognition of the possibility that *no agent* (playing the role he assumes) is really necessary, even in a telepathic experience.

If the first requirement of an adequate hypothesis is that it cover all the facts, the second is that it be testable. Even if the hypothesis under discussion had fulfilled the first requirement, could it have been subjected to test? It would be necessary first of all to bring to proof the hypothesis of mind on which it rests. As already discussed, the handicap of the old concept of telepathy lies in the still unknown relationship between thought and brain. Neither the early concept of mind-to-mind transfer nor Tyrrell's idea of "non-separate mid-level" constituents of the unconscious has been supported by any experimental evidence beyond the fact that the essential psi operation is known to be unconscious. Up to the present all these theories remain untested, and thus their ultimate value is indeterminate.

This book is testimony to Tyrrell's broad vision of the significance of spontaneous psi experiences for the larger understanding of the personality of man. Every serious insightful attempt to evaluate such exceptional manifestations as those with which the study deals has also an important educational function.

Louisa E. Rhine

PSYCHICAL RESEARCH TODAY. By Donald J. West. London: Gerald Duckworth and Co., Ltd., 1954. 144 pp. 12*s* 6*d*.

This moderately priced little book is probably the best at present available to put into the hands of an intelligent inquirer who wants up-to-date information about psychical research. It is interestingly written, well-informed, and, while not neglecting the experimental side of the work, it puts it in relation to spontaneous and mediumistic phenomena. It takes account of experimental work done on both sides of the Atlantic and gives clairvoyance due weight as well as telepathy. It is true that only two pages are devoted to psychokinesis and that the only experiment specifically referred to is a recent one by Mr. Fisk; but that is two more pages than some psychical researchers would allow to psychokinesis. ESP researches are more fully recorded and include accounts of work at the Duke Laboratory as well as that of Soal, Tyrrell, and Carington. Recent work in ESP is also mentioned, including that of Mr. Fisk and the experiments of Dr. Osis with cats.

There are interesting accounts of spontaneous phenomena and a somewhat negative reference to haunted houses and poltergeists.

The case of Eva C. is discussed as an example of a physical medium, and the final chapter deals with some of the theories that have been put forward, including Whately Carington's association theory and Professor Hardy's theory of telepathy in evolution.

Although Dr. West's book is interesting and informative, it can hardly be recommended as a guide to the uninformed inquirer without some reservations. The author's judgment is less to be relied on than his information, and his bias is apparent both in his treatment of controversial questions and in his selection of evidence. It is true that no book on psychical research that is not to be of intolerable length can be written without selection and we should not all agree as to what cases should be selected. But it is necessary that selection should not be in the direction of consistent under-estimation or over-estimation of the strength of evidence in controversial matters. Dr. West's selections seem generally to be in the direction of under-estimating the evidence. There are, for example, cases of alleged physical mediumship more difficult to dismiss than that of Eva C., such as the less spectacular but more carefully observed phenomena of Anna Rasmussen, which are not mentioned. One may also doubt whether a fair picture of the evidence for poltergeists is given when there is no mention of Barrett's study of the poltergeist phenomena of Derrygonelly Farm, which is much stronger evidence than any mentioned in the book.

Dr. West gives the impression that generally through the book he is trying to be fair even to those with whom he disagrees. There seem, however, to be some lapses from this standard. On page 114, for example, there is an unfortunately worded reference to Dr. Rhine. On page 52, the value of the testimony of Crookes to the phenomena of D. D. Home is questioned on the ground that "he all but flirted with Katie King, a realistic feminine spirit materialized by a medium called Florence Cook, who was repeatedly exposed in fraud." Professor Richet is said to be "credulous as regards physical phenomena." This is making the matter unfairly easy for the critic. The difficult thing to explain about the Eva C. phenomena is why, in spite of much that seems dubious to us, they seemed impressive to a scientist so well qualified to judge as was Richet. The difficulty is real and not to be got over by sneering at Richet as credulous.

There are also some errors of fact. For example, on page 48 it is stated that the séances of Eva C. were held in darkness. The report by Richet in *Thirty Years of Psychical Research* says, on the contrary: "Light was given by a photographer's red lamp. . . . Everything that took place in the room could be seen perfectly well." It is difficult to understand how a figure draped in white could have been seen (as Dr. West reports in the same sentence) if the room had been dark.

R. H. THOULESS

GLOSSARY

In order to avoid constant redefining of commonly recurring terms in papers appearing in this JOURNAL, the following definitions are submitted for convenient reference.

A simple description of the basic experimental methods and of the statistical procedures to be followed in evaluating ESP (both card tests and drawings tests) and PK may be found in the *Handbook of Tests in Parapsychology* by Betty M. Humphrey (published by the Parapsychology Laboratory, Duke University, Durham, N. C.).

AGENT: In tests for telepathy, the person whose mental states are to be apprehended by the percipient. In GESP tests, the person who looks at the stimulus object.

AVERAGE SCORE: Average number of hits per run.

BT ("Broken Technique"): The clairvoyance technique in which each card is laid aside by the experimenter as it is called by the subject. The check-up is made at the end of the run.

CALL: The ESP symbol selected by the subject in trying to identify a target.

CHANCE: The complex of undefined causal factors irrelevant to the purpose at hand.

CHANCE EXPECTATION = MEAN CHANCE EXPECTATION: The most likely score if only chance obtains.

CHANCE AVERAGE: Mean chance expectation expressed as an expected score, generally in terms of average per run.

CHI SQUARE: A sum of quantities each of which is a deviation squared divided by an expected value. Also a sum of the squares of CR's.

CLAIRVOYANCE: Extrasensory perception of objective events as distinguished from telepathic perception of the mental state of another person.

CR (CRITICAL RATIO): A measure to determine whether or not the observed deviation is significantly greater than the expected random fluctuation about the average. The CR is obtained by dividing the observed deviation by the standard deviation. (The probability of a given CR may be obtained by consulting tables of the probability integral, such as Pearson's.)

CR OF THE DIFFERENCE: The observed difference between the average scores of two samples of data divided by the standard deviation of the difference.

DECK: Twenty-five ESP cards.

CLOSED DECK: An ESP deck composed of five each of the five symbols.

OPEN DECK: An ESP deck made up of the ESP symbols selected in random order.

DEVIATION: The amount an observed number of hits or an average score varies from the mean chance expectation or chance average. A deviation may be total (for a series of runs) or average (per run).

DIE THROW: The elementary unit of data in a dice-throwing test represented by the throwing and reading of a single die, no matter how many dice are thrown at a time.

DISPLACEMENT: ESP responses to targets other than those for which the calls were intended.

BACKWARD DISPLACEMENT: ESP responses to targets preceding the assigned targets (the ones for which they were intended). Displacement to the targets

one, two, three, etc. places preceding the assigned target are designated as (−1), (−2), (−3), etc.

FORWARD DISPLACEMENT: ESP responses to targets coming later than the assigned targets. Displacement to the targets one, two, three, etc. places after the assigned target are designated as (+1), (+2), (+3), etc.

DT (Down Through): The clairvoyance technique in which the cards are called down through the deck before any are removed or checked.

EMPIRICAL CONTROL: An experiment which wholly or partially follows the main experiment with the exception that the conditions are designed to exclude the possibilty of the hypothesis being tested.

ESP (EXTRASENSORY PERCEPTION): Response to an external event not presented to any known sense.

ESP CARDS: Cards, each bearing one of the following five symbols: star, circle, three parallel wavy lines (called "waves"), square, plus.

ESP SYMBOLS: See plate opposite page 1, this JOURNAL, Vol. 1, March, 1937.

EXPECTATION; see CHANCE.

EXTRACHANCE: Not due to chance alone.

FREE MATERIAL: The stimulus objects of experiments in which an unlimited or unspecified range of stimulus objects is employed (as contrasted with methods such as card-calling in which the subject knows that the stimulus object is one of a known range).

FREE RESPONSE METHOD: An ESP test method in which a relatively unlimited range of stimulus objects is used and in which the subject is instructed to express freely as his response any range of thoughts and ideas.

GESP (GENERAL EXTRASENSORY PERCEPTION): A technique designed to test the occurrence of extrasensory perception, permitting either telepathy or clairvoyance or both to operate.

HIGH-DICE TEST: A PK technique in which the aim of the subject is to try to influence a pair of dice to fall with the two upper faces totaling eight or more.

LOW-DICE TEST: A PK technique in which the aim of the subject is to try to influence a pair of dice to fall with the two upper faces totaling six or less.

MEAN CHANCE EXPECTATION; see CHANCE.

P (PROBABILITY): A mathematical estimate of the expected relative frequency of a given event if chance alone were operative.

PARAPSYCHOLOGY: A division of psychology dealing with those psychical effects which appear not to fall within the scope of what is at present recognized law.

PERCIPIENT: The person who makes the calls in an ESP test.

PK (PSYCHOKINESIS): The direct influence exerted on a physical system by a subject without any known intermediate physical energy or instrumentation.

PRECOGNITION: Cognition of a future event which could not be known through rational inference.

PREFERENTIAL MATCHING: A method of scoring free responses. A judge ranks the stimulus objects (usually in sets of four) with respect to their similarity to, or association with, each response; and/or he ranks the responses with respect to their similarity to, or association with, each stimulus object.

PSI: A general term to identify personal factors or processes in nature which transcend accepted laws. It approximates the popular use of the word "psychic" and the technical one, "parapsychical."

PSI PHENOMENA: Occurrences which result from the operation of psi. They include the phenomena of both ESP (including precognition) and PK.

QD (QUARTER DISTRIBUTION): The distribution of hits in the record page (or in a logical subdivision thereof, such as the set or the half-set) as found in the four equal quarters formed by dividing the selected unit horizontally and vertically.

RUN: A group of trials, usually the successive calling of a deck of 25 ESP cards or symbols. In PK tests, 24 single die throws regardless of the number of dice thrown at the same time.

SCORE: The number of hits made in one run.

TOTAL SCORE: Total of scores made in a given number of runs.

AVERAGE SCORE: Total score divided by number of runs.

SD (STANDARD DEVIATION): Usually the theoretical root mean square of the deviations. It is obtained from the formula $\sqrt{npq}$, in which n is the number of single trials, p the probability of success per trial, and q the probability of failure. (For ESP cards, SD $= 2\sqrt{\text{no. of runs.}}$) Sometimes the SD used is that estimated from the observed variability in the scores.

SD OF THE DIFFERENCE: For both ESP cards and PK tests using dice, the SD of the difference is equal to $\sigma_s \sqrt{1/R_1 + 1/R_2}$ where σ^s is the SD of a single run and R_1 and R_2 are the number of the runs in the respective samples compared. This gives the SD of the difference for run score averages.

SERIES: Several runs or experimental sessions that are grouped in accordance with a stated principle.

SESSION: A unit of an ESP or PK experiment comprising all the trials of one test occasion.

SET: A subdivision of the record page serving as a scoring unit for a consecutive group of trials, usually for the same target.

SEVENS TEST: A PK technique in which the aim of the subject is to try to influence a pair of dice to fall with the two upper faces totaling seven.

SIGNIFICANCE: A numerical result is significant when it equals or surpasses some criterion of degree of chance improbability. The criterion commonly used in this JOURNAL is a probability value of .02 or less, or a deviation in either direction such that the CR is 2.33 or greater.

SINGLES TEST: A PK technique in which the aim of the subject is to try to influence dice to fall with a specified face up.

STIMULUS OBJECT: The ESP card or drawing or other object, some identifying characteristic of which is to be apprehended by the percipient.

SUBJECT: The person who is experimented upon. In ESP tests, most commonly the percipient (though also the agent in GESP and telepathy). In PK tests, any individual whose task it is to influence the objects thrown.

TARGET: In ESP tests, the stimulus object; or in telepathy, the mental state of the agent. In PK tests, the faces of the die (or combination of faces) which the subject attempts to bring up in the act of throwing.

TARGET CARD: The card which the percipient is attempting to perceive (i.e., to identify or otherwise indicate a knowledge of).

TARGET DECK: The deck of cards the order of which the subject is attempting to identify.

TARGET FACE: The face on the die which the subject tries to turn up as a consequence of direct mental action.

TELEPATHY: Extrasensory perception of the mental activities of another person. It does not include the clairvoyant perception of objective events.

TRIAL: In ESP tests, a single attempt to identify a stimulus object. In PK tests, a single throw of the dice or other objects thrown.

radio signals, we can calculate that they would fall below noise level within a few millimeters from the surface of the head.

"Even if we ignore these physical characteristics, the observations reported on extrasensory phenomena seem to exclude any such approach; for there is no evidence that screening of the subject, or distance between sender and receiver, has any influence on the nature or abundance of the effects described. Furthermore, it seems to be one of the cardinal claims of workers in this field that a signal may be received before it is transmitted. If we accept these observations for what they are said to be, we cannot fit them into the physical laws of the universe as we define them today. We may reject the claims of transcendental communication on the grounds of experimental error or statistical fallacy, or we may withhold judgment, or we may accept them gladly as evidence of spiritual life: but it does not seem easy to explain them in terms of biological mechanism."

W. Grey Walter, *The Living Brain**

* Published by W. W. Norton and Co.

HERE AND THERE IN THE LITERATURE

"We must confess at this stage that no study of brain activity has thrown any light on the peculiar forms of behavior known variously as secondsight, clairvoyance, telepathy, extrasensory perception and psychokinesis. It has often been suggested by those seeking a material basis for otherwise unaccountable behavior that the electrical activity of the brain might be the mechanism whereby information could be transmitted from brain to brain, and that the electrical sensitivity of the brain might be a means of communicating with some all-pervading influence. Quite apart from any philosophic objection there may be to such an argument, the actual scale and properties of the brain's electrical mechanisms offer no support for it. The size of the electrical disturbances which the brain creates are extremely small. In fact, they are about the size, within the brain itself, of a received signal which is just intelligible on an average radio set. More crucial even than this, their dominant frequencies are far below the range of radio channels, below even the scale of audible frequencies. At ten cycles per second, the average frequency of the alpha rhythms, any electromagnetic signal transmitted through space would have a wave length of thirty million meters.

"The familiarity of radio signalling around the world has popularized the notion that any signal once generated may be propagated indefinitely through the chasms of space, so that all events have an eternal quality in some attenuated but identifiable form. This is not even approximately true; for any signal, however propagated, weakens with its passage until its size falls below the level of noise and interference in some locality. Beyond this point it can never be detected, however great the resolution and selectivity of the receiver. If we consider the largest rhythms of the brain as casual

(Continued on third cover)